PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6: C12N 15/53, 9/02, A61K 39/395, C12Q 1/68, G01N 33/577 // C12P 21/02 — A1

(11) International Publication Number: **WO 96/12810**

(43) International Publication Date: 2 May 1996 (02.05.96)

(21) International Application Number: PCT/GB95/02465

(22) International Filing Date: 18 October 1995 (18.10.95)

(30) Priority Data:
9421093.7 19 October 1994 (19.10.94) GB

(71) Applicant *(for all designated States except US)*: THE UNIVERSITY OF EDINBURGH [GB/GB]; Old College, South Bridge, Edinburgh EH8 9YL (GB).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: LATHE, Richard [GB/GB]; The University of Edinburgh, Centre for Genome Research, King's Buildings, West Main Road, Edinburgh EH9 3JQ (GB). ROSE, Kenneth, Andrew [GB/GB]; The University of Edinburgh, Centre for Genome Research, King's Buildings, West Mains Road, Edinburgh EH9 3JQ (GB). STAPLETON, Genevieve [AU/GB]; The University of Edinburgh, Centre for Genome Research, King's Buildings, West Mains Road, Edinburgh EH9 3JQ (GB).

(74) Agent: RITTER, Stephen, David; Mathys & Squire, 100 Grays Inn Road, London WC1X 8AL (GB).

(81) Designated States: AL, AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TT, UA, UG, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: HIPPOCAMPUS-ASSOCIATED PROTEINS, DNA SEQUENCES CODING THEREFOR AND USES THEREOF

(57) Abstract

This invention provides novel hippocampus-associated proteins and DNA sequences coding therefor. In an investigation of hippocampus-associated proteins by differential screening of a rat hippocampus cDNA library, a cDNA species encoding a novel protein designated Hct-1 was isolated and shown to be related to cytochromes P450. The use of hybridization probes based on the rat Hct-1 sequence has led to the identification of homologues in other mammalian species.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GB**	United Kingdom	**MR**	Mauritania
AU	Australia	**GE**	Georgia	**MW**	Malawi
BB	Barbados	**GN**	Guinea	**NE**	Niger
BE	Belgium	**GR**	Greece	**NL**	Netherlands
BF	Burkina Faso	**HU**	Hungary	**NO**	Norway
BG	Bulgaria	**IE**	Ireland	**NZ**	New Zealand
BJ	Benin	**IT**	Italy	**PL**	Poland
BR	Brazil	**JP**	Japan	**PT**	Portugal
BY	Belarus	**KE**	Kenya	**RO**	Romania
CA	Canada	**KG**	Kyrgystan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SI**	Slovenia
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SK**	Slovakia
CM	Cameroon	**LI**	Liechtenstein	**SN**	Senegal
CN	China	**LK**	Sri Lanka	**TD**	Chad
CS	Czechoslovakia	**LU**	Luxembourg	**TG**	Togo
CZ	Czech Republic	**LV**	Latvia	**TJ**	Tajikistan
DE	Germany	**MC**	Monaco	**TT**	Trinidad and Tobago
DK	Denmark	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**US**	United States of America
FI	Finland	**ML**	Mali	**UZ**	Uzbekistan
FR	France	**MN**	Mongolia	**VN**	Viet Nam
GA	Gabon				

HIPPOCAMPUS-ASSOCIATED PROTEINS, DNA SEQUENCES CODING THEREFOR AND USES THEREOF

This invention relates to novel hippocampus-associated proteins, to DNA sequences coding therefor, to uses thereof and to antibodies to said proteins. The novel hippocampus-associated proteins are believed to be of the cytochrome P450 class.

Background to the Invention

The identification of hippocampus-associated proteins and the isolation of cDNA molecules coding therefor is important in the field of neurophysiology. Thus, for example, such proteins are believed to be associated with memory functions and abnormalities in these proteins, including abnormal levels of expression and the formation of modified or mutated protein is considered to be associated with pathological conditions associated with memory impairment. The isolation of novel hippocampus-associated proteins and the associated DNA sequences coding therefor is consequently of considerable importance.

The present invention arose out of our investigation of hippocampus-associated proteins by differential screening of a rat hippocampus cDNA library. A cDNA species encoding a novel protein which we have designated Hct-1 was isolated and shown to be related to cytochromes of the P450 class.

The use of hybridization probes based on the rat Hct-1 sequence has led to the identification of homologues in other mammalian species, specifically mouse and human.

Cytochromes P450 are a diverse group of heme-containing mono-oxygenases (termed CYP's; see Nelson *et al.*, DNA Cell Biol. (1993) **12**, 1-51) that catalyse a variety of oxidative conversions, notably of steroids but also of fatty acids and xenobiotics. While CYP's are most abundantly expressed in the testis, ovary, placenta, adrenal and liver, it is becoming clear that the brain is a further site of

CYP expression. Several CYP activities or mRNA's have been reported in the nervous system but these are predominantly of types metabolizing fatty acids and xenobiotics (subclasses CYP2C, 2D, 2E and 4). However, primary rat brain-derived glial cells have the capacity to synthesize pregnenolone and progesterone *in vitro*. Mellon and Deschepper, Brain Res. (1993), **629**, 283-292(9) provided molecular evidence for the presence, in brain, of key steroidogenic enzymes CYP11A1 (scc) and CYP11B1 (11ß) but failed to detect CYP17 (c17) or CYP11B2 (AS). Although CYP21A1 (c21) activity is reported to be present in brain, authentic CYP21A1 transcripts were not detected in this tissue.

Interest in steroid metabolism in brain has been fuelled by the finding that adrenal- and brain-derived steroids (neurosteroids) can modulate cognitive function and synaptic plasticity. For instance, pregnenolone and steroids derived from it are reported to have memory enhancing effects in mice. However, the full spectrum of steroid metabolizing CYP's in brain and the biological roles of their metabolites *in vivo* has not been established.

To investigate such regulation of brain function our studies have focused on the hippocampus, a brain region important in learning and memory. Patients with lesions that include the hippocampus display pronounced deficits in the acquisition of new explicit memories while material encoded long prior to lesion can still be accessed normally. In rat, neurotoxic lesions to the hippocampus lead to a pronounced inability to learn a spatial navigation task, such as the water maze. The role of the hippocampus in learning has been further emphasized by the finding that hippocampal synapses, notably those in region CA1, display a particularly robust form of activity-dependent plasticity known as long term potentiation (LTP). This phenomenon satisfies some of the requirements for a molecular mechanism underlying memory processes - persistence, synapse-specificity and associativity. LTP is thought to be initiated by calcium influx through the NMDA (N-methyl D-aspartate) subclass of receptor activated by the excitatory neurotransmitter, L-glutamate, and occlusion of NMDA receptors *in vivo* with the competitive

antagonist AP5 both blocks LTP and the acquisition of the spatial navigation task.

The induction of LTP is attenuated by simultaneous release of gamma-amino butyric acid (GABA) from inhibitory interneurons: activation of $GABA_A$ receptors antagonizes L-glutamate induced depolarization of the postsynaptic neuron and interplay between the GABA and L-glutamate receptor pathways is thought to modulate the establishment of LTP. Interplay between these two circuits is emphasised by the finding that some aesthetics (e.g. ketamine) act as antagonists of the NMDA receptor while others, such as the steroid aesthetic alfaxolone, are thought to be agonists of the $GABA_A$ receptor. It is of particular note that some naturally occurring steroids, such as pregnenolone sulfate, act as agonists of the $GABA_A$ receptor, while pregnenolone sulfate is also reported to increase NMDA currents. Although neurosteroids principally appear to exert their effects via the $GABA_A$ and NMDA receptors, there have been indications that neurosteroids may also interact with sigma and progesterone receptors.

Despite considerable interest in the action of neuro-active steroids, and possible roles in modulating synaptic plasticity and brain function, little is known of pathways of steroid metabolism in the central nervous system. As part of a study into the molecular biology of the hippocampal formation, and the mechanisms underlying synaptic plasticity, we have sought molecuiar clones corresponding to mRNA's expressed selectively in the formation. One such cDNA, Hct-1 (for *h*ippo*c*ampal *t*ranscript), was isolated from a cDNA library prepared from adult rat hippocampus. Sequence analysis has revealed that Hct-1 is a novel cytochrome P450 most closely related to cholesterol- and steroid-metabolizing CYP's but, unlike other CYP's, is predominantly expressed in brain. The present invention provides molecular characterization of Hct-1 coding sequences from rat, mouse and humans, their expression patterns, and discusses the possible role of Hct-1 in steroid metabolism in the central nervous system.

DNA sequences encoding hitherto unknown cytochrome P450 proteins have now been identified and form one aspect of the present invention.

Summary of the Invention

According to one aspect of the present invention there are thus provided DNA molecules selected from the following:

(a) DNA molecules containing the coding sequence set forth in SEQ Id No: 1 beginning at nucleotide 22 and ending at nucleotide 1541,

(b) DNA molecules containing the coding sequence set forth in SEQ Id No: 2 beginning at nucleotide 1 and ending at nucleotide 1242,

(c) DNA molecules capable of hybridizing with the DNA molecule defined in (a) or (b) under standard hybridization conditions defined as 2 x SSC at 65°C.

(d) cytochrome P450-encoding DNA molecules capable of hybridizing with the DNA molecule defined in (a), (b) or (c) under reduced stringency hybridization conditions defined as 6 x SSC at 55°C.

Such DNA sequences can represent coding sequences of Hct-1 proteins. The sequences (a) and (b) above represent the mouse and rat Hct-1 gene sequence. Homologous sequences from other vertebrate species, especially mammalian species (including man) fall within the class of DNA molecules represented by (c) or (d).

Thus the present invention further provides a DNA molecule consisting of sequences of the human Hct-1 gene.

These DNA sequences may be selected from the following:

(e) DNA molecules comprising one or more sequences selected from

(i) the sequence designated "intron 2" in SEQ Id No 3,

(ii) the sequence designated "exon 3" in SEQ Id No 3,

(iii) the sequence designated "intron 3" in SEQ Id No 3,

(iv) the sequence designated "exon 4" in SEQ Id No 3, and

(v) the sequence designated "intron 5" in SEQ Id No 3; and

(f) DNA molecules capable of hybridizing with the DNA molecules defined in (e) under standard hybridization conditions defined as 2 x SSC at 65°C.

(g) cytochrome P450-encoding DNA molecules capable of hybridizing with the DNA molecule defined in (e) or (f) under reduced stringency hybridization conditions defined as 6 x SSC at 55°C.

(h) DNA molecules comprising contiguous pairs of sequences selected from

(i) the sequence designated "intron 2" in SEQ Id No 3,

(ii) the sequence designated "exon 3" in SEQ Id No 3,

(iii) the sequence designated "intron 3" in SEQ Id No 3,

(iv) the sequence designated "exon 4" in SEQ Id No 3, and

(v) the sequence designated "intron 5" in SEQ Id No 3; and

(i) DNA molecules capable of hybridizing with the DNA molecules defined in (h) under standard hybridization conditions defined as 2 x SSC at 65°C.

(j) cytochrome P450-encoding DNA molecules capable of hybridizing with the DNA molecule defined in (h) or (i) under reduced stringency hybridization conditions defined as 6 x SSC at 55°C.

(k) DNA molecules comprising a contiguous coding sequence consisting of the sequences "exon 3" and "exon 4" in SEQ Id No 3, and

(l) DNA molecules capable of hybridizing with the DNA molecules defined in (k) under standard hybridization conditions defined as 2 x SSC at 65°C.

(m) cytochrome P450-encoding DNA molecules capable of hybridizing with the DNA molecule defined in (k) or (l) under reduced stringency hybridization conditions defined as 6 x SSC at 55°C.

It will be appreciated that the DNA sequences that include introns (such as the sequences covered by definitions (e) to (j) above), may consist of or be derived from genomic DNA. Those sequences that exclude introns may also be genomic in origin, but typically would consist of or be or be derived from cDNA. Such sequences could be obtained by probing an appropriate library (cDNA or genomic) using hybridisation probes based upon the sequences provided according to the invention, or they could be prepared by chemical synthesis or by ligation of sub-sequences.

The invention further provides DNA molecules encoding an Hct-1 gene-associated sequence coded for by a DNA molecule as defined above, but which differ in sequence from said sequences by virtue of one or more amino acids of said Hct-1 gene-associated sequences being encoded by degenerate codons.

The present invention further provide DNA molecules useful as hybridization probes and consisting of a contiguous sequence of at least 18 nucleotides from the DNA sequence set forth in SEQ Id Nos: 1, 2 and 3.

Such molecules preferably contain at least 24 and more preferably at least 30 nucleotide taken from said sequences.

The aforementioned DNA molecules are useful as hybridization probes for isolating members of gene families and homologous DNA sequences from different species. Thus, for example, a DNA sequence isolated from one rodent species, for example rat, has been used for isolating homologous sequences from another rodent species, for example mouse and from other mammalian species , e.g. primate species such as humans.

Such sequences may be further used for isolating homologous sequences from other mammalian species, for example domestic animals such as cows, horses, sheep and pigs, primates such as chimpanzees, baboons and gibbons.

DNA sequences according to the invention may be used in diagnosis of neuropsychiatric disorders, endocrine disorders, immunological disorders, diseases of cognitive function, neurodegenerative diseases or diseases of cognitive function, for example by assessing the presence of depleted levels of mRNA and/or the presence of mutant or modified DNA molecules. Such sequences include hybridisation probes and PCR primers. The latter generally would be short (e.g. 10 to 25) oligonucleotides in length and would be, capable of hybridising with a DNA molecule as defined above. The invention includes the use of such primers in the detection of genomic or cDNA from a biological sample for the purpose of diagnosis of neuropsychiatric disorders, endocrine disorders, immunological disorders, diseases of cognitive function or neurodegenerative diseases.

The present invention further provides hippocampus-associated proteins as such, encoded by the DNA molecules of the invention.

In particular, there is provided

(i) the protein designated rat Hct-1 comprising the amino acid sequence set forth in SEQ Id No: 1 or a protein having substantial homology thereto,

(ii) the protein designated mouse Hct-1 comprising the amino acid sequence set forth in SEQ Id No: 2 or a protein having substantial homology thereto, or

(iii) the protein designated human Hct-1 comprising the amino acid sequence set forth in SEQ Id No: 3 or a protein having substantial homology thereto.

By "substantial homology" is meant a degree of homology such that at least 50%, preferably at least 60% and most preferably at least 70% of the amino acids match. The invention of course covers related proteins having a higher degree of homology, e.g. at least 80%, at least 90% or more.

The Hct-1 polypeptides may be produced in accordance with the invention by culturing a transformed host and recovering the desired Hct-1 polypeptide, characterised in that the host is transformed with nucleic acid comprising a coding sequence as defined above.

Examples of suitable hosts include yeast, bacterial, insect or mammalian cells. Although vectorless expression may be employed, it is preferred that the nucleic acid used to effect the transformation comprises an expression construct or an expression vector, e.g. a vaccinia virus, a baculovirus vector, a yeast plasmid or integration vector.

The invention further provides antibodies, especially monoclonal antibodies which bind to Hct-1 proteins. These and the proteins of the invention may be employed in the design and/or manufacture of an antagonist to Hct-1 protein for diagnosis and/or treatment of diseases of cognitive function or neurodegenerate diseases. The use of Hct-1-associated promoters in the formation of constructs for use in the creation of transgenic animals is also envisaged according to the invention. The antibodies of the invention may be prepared in conventional manner, i.e. by immunising animal such as rodents or rabbits with purified protein obtained from recombinant yeast, or by immunising with recombinant vaccinia.

Hct-1 proteins provided according to the invention posseses catalytic activity, thus they may be used in industrial processes, to effect a catalytic transformation of a substrate. For example, where the substrate is a steroid, the proteins may be used to catalyse stereospecific transformations, e.g. transformations involving oxygen transfer.

Description of Drawings (see also Figure legends - 7 *infra*)

Figure 1 illustrates (a) a restriction map of clone 12 and (b) the complete nucleotide and translation sequence of the 1.4kb cDNA clone of rat Hct-1,

Figure 2 illustrates Northern analysis of Hct-1 expression in adult rat and mouse brain, and other tissues,

Figure 3 illustrates (a) restriction maps of clones 35 and 40 and (b) the complete nucleotide and translation sequence of mouse Hct-1 cDNA,

Figure 4 illustrates an alignment of mouse Hct-1 with human CYP7 and highlights regions homologous to other steroidogenic P450s,

Figure 5 illustrates an analysis of Hct-1 expression in mouse brain,

Figure 6 illustrates Southern analysis of Hct-1 coding sequences in mouse, rat and human.

Figure 7 illustrates Southern blot analyses of mouse genomic DNA using (a) a full length mouse Hct-1cDNA clone and (b) rat genomic DNA probed with clone 14.5a,

Figure 8 illustrates a genomic map of mouse Hct-1,

Figure 9 illustrates a partial nucleotide sequence of human genomic Hct-1 (CYP7B1) and the encoded polypeptide,

Figure 10 illustrates an amino acid alignment of mouse Hct-1 and human CYP7,

Figure 11A illustrates Kozak sequences in mRNAs for steroidogenic P540's,

Figure 11B illustrates mutagenesis of the 5'end of the mouse Hct-1 cDNA to sreate a near-consensus translation initiation region surrounding the ATG (AUG),

Figure 12 illustrates yeast expression vectors containing the mouse Hct-1 coding sequence, and

Figure 13 illustrates a vaccinia expression vectors containing the mouse Hct-1 coding sequence.

Description of Specific Embodiments

Details of the isolation of hippocampus-associated DNA molecules according to the invention will now be described by way of example:

1. ISOLATION OF GENE ENCODING RAT HCT-1

1.1 Differential screening of a rat hippocampus cDNA library

To identify genes whose expression is enriched in the hippocampal formation we performed a differential hybridization screen of a hippocampal cDNA library. Adult rat hippocampal RNA was reverse transcribed using a oligo-dT-NotI primer, converted to double-stranded cDNA, EcoRI adaptors were attached and the cDNA's were inserted between the EcoRI and NotI sites of a bacteriophage lamda vector.

1.1.1 Preparation of cDNA libraries

Following anaesthesia (sodium pentobarbital) of adult rats (Lister hooded) the hippocampal formation was dissected, including areas CA1-3 and dentate gyrus,

subiculum, alvear and fimbrial fibres but excluding fornix and afferent structures such as septum and entorhinal cortex. Remainder of brain was also pooled taking care to exclude hippocampal tissue. Total RNAs were prepared by a standard guanidinium isothiocyanate procedure, centrifugation through a CsCl cushion, and poly-A^+ mRNA selected by affinity chromatography on oligo-dT cellulose. First strand cDNA synthesis used a NotI adaptor primer

[5-dCAATTCGCGGCCGC$(T)_{15}$-3']

and Moloney murine leukemia virus (MMLV) reverse transcriptase; second strand synthesis was performed by RNaseH treatment, DNA polymerase I fill-in and ligase treatment. Following the addition of hemi-phosphorylated EcoRI adaptors (5'-dCGACAGCAACGG-3' and 5'-dAATTCCGTTGCTGTCG-3') and cleavage with NotI the cDNA was inserted between the NotI and EcoRI sites of bacteriophage lambda vector lambda-ZAPII (Stratagene).

1.1.2 Differential hybridization screening

Recombinant bacteriophage plaques were transferred in duplicate to Hybond-N membranes (Amersham), denatured (0.5 M NaOH, 1.5 M NaCl, 4 min), renatured (1 M Tris.HCl pH 7.4, 1.5 M NaCl), rinsed, dried and baked (2 h, 80°C). Hybridization as described (Church et al., Proc. Natl. Acad. Sci. USA (1984), **81** 1991-1995) used a radiolabelled probe prepared by MMLV reverse transcriptase copying of polyA$^+$ RNA (from either hippocampus or the remainder of brain) into cDNA in the presence of α-^{32}P-dCTP and unlabelled dGTP, dATP and dTTP according to standard procedures. Following washing and exposure for autoradiography, differentially hybridizing plaques were repurified. Inserts were transferred to a pBluescript vector either by cleavage and ligation or by using in vivo excision using the ExAssist/SOLR system (Stratagene).

Duplicate lifts from 500,000 plaques were screened with radiolabelled cDNA probes prepared by reverse transcription of RNA from either hippocampus (Hi) or 'rest of brain' (RB). Approximately 360 clones gave a substantially stronger hybridization signal with the Hi probe than with the RB probe; 49 were analysed

in more depth. *In vivo* excision was used to transfer the inserts to a plasmid vector for partial DNA sequence studies. Of these, 21 were novel (not presented here); others were known genes whose expression is enriched in hippocampus but not specific to the formation (eg., the rat amyloidogenic protein. Northern analysis was first performed using radiolabelled probes corresponding to the 21 novel sequences. While three (12.10a, 14.5a and 15.13a) identified transcripts specific to the hippocampus, 12.10a and 15.13a both hybridized to additional transcripts whose expression was not restricted to the formation. Clone 14.5a appeared to identify transcripts enriched in hippocampus and was dubbed Hct-1.

1.2 Characterisation of Rat Hct-1

1.2.1 Rat Hct-1 encodes a cytochrome P450

To extend this characterization, the insert of clone 14.5a (300 nt) was used to rescreen the hippocampal cDNA library. 4 positives were identified (clones 14.5a-5, -7, -12 and -13), and the region adjacent to the poly-A tail analysed by DNA sequencing. While clones 5 (0.7 kb) and 12 (1.4 kb) had the same 3' end as the parental clone, clone 7 (0.9 kb) had a different 3' end consistent with utilization of an alternative polyadenylation site. Clone 13 (2.5 kb), however, appeared unrelated to Hct-1 and was dubbed Hct-2.

Clones 12 and 7 were then fully sequenced and the sequences obtained were compared with the database. Significant homology was detected between clone 12 and the human and rat cDNA's encoding cholesterol 7α-hydroxylase, though the sequences are clearly distinct. At the nucleic acid level, the 1428 nt cDNA clone for rat Hct-1 shared 55% identity over an 1100 nt overlap with human cholesterol 7α-hydroxylase (CYP7) and 54% identity over a 1117 nt overlap with rat CYP7. Fig. 1 gives the partial cDNA sequences of rat Hct-1 and the encoded polypeptide.

1.2.2 Hct-1 mRNA expression in rat

Rat Hct-1 clone 14.5a/12 (1.4 kb) was used to investigate the expression of Hct-1 mRNA in rat brain and other organs. We first performed in situ hybridization to sections of rat brain. While these preliminary experiments did not permit unambiguous localization of Hct-1 transcripts, we confirmed expression in the hippocampus, predominantly in the cell layers of the dentate gyrus, while weaker expression was detected in other hippocampal and brain regions (not presented). Northern analysis was then performed on RNA prepared from different sections of rat brain. In Fig. 2*A* the Hct-1 probe identifies three transcripts in hippocampus of 5.0, 2.1 and 1.8 kb, with the two smaller transcripts being particularly enriched in hippocampus. The larger transcript was only detectable in brain, while the two smaller transcripts were also present in liver (and, at much lower levels, in kidney) but not in other organs tested including adrenal (not shown), testis, and ovary. In brain, expression was also detected in olfactory bulb and cortex while very low levels were present in cerebellum (Fig. 2*A*).

1.2.3 Sexual dimorphism of Hct-1 expression in liver but not in brain

The expression of several CYPS is known to be sexually dimorphic in liver. We therefore inspected liver and brain of male and female rats for the presence of Hct-1 transcripts. In Fig. 2*B* the Hct-1 probe revealed the 1.8 and 2.1 kb (and 5.0 kb, Hct-2) transcripts in both male and female brain, with the 2.1 kb Hct-1 transcript predominating. Levels of Hct-1 mRNA's in liver were reduced greater than 20-fold over those detected in brain. Furthermore, Hct-1 transcripts were only significant in liver from male animals; expression of Hct-1 in females was barely detectable demonstrating that hepatic expression of Hct-1 is sexually dimorphic.

2. ISOLATION OF MOUSE HCT-1

2.1 Isolation of mouse Hct-1 cDNA clones

A mouse liver cDNA library, established as NotI-EcoRI fragments in a lambda-gt10 vector, was probed using a rat Hct-1 probe. The library was a kind gift of B. Luckow and K. Kästner, Heidelberg.

Because the transcripts identified by the Hct-1 probe (predominantly 1.8 and 2.1 kb) are clearly longer than the longest cDNA clone (1.4 kb) obtained from our rat hippocampus library, we therefore elected to pursue studies with the mouse Hct-1 ortholog. A mouse liver cDNA library was screened using a rat Hct-1 probe and four clones were selected, none containing a poly-A tail. Two (clones 33 and 35, both 1.8 kb) gave identical DNA sequences at both their 5' and 3' ends, and this sequence was approximately *91%* similar to rat Hct-1. The remaining two clones, 23 and 40, were also identical to each other and were related to the other clones except for a 5' extension in (59 nt) and a 3' deletion (99 nt). The complete DNA sequences of clones 35 and 40 were therefore determined.

The sequences obtained were identical throughout the region of overlap. The mouse Hct-1 open reading frame (ORF) commences with a methionine at nucleotide 81 (numbering from clone 40) and terminates with a TGA codon at nucleotide 1600, encoding a protein of 507 amino acids (Fig. 3). At the 5' end it is of note that the ATG initiation codon leading the ORF does not correspond to the translation initiation consensus sequence YYAYYATGR. However, the 5' untranslated region cloned is devoid of other possible initiation codons and an in-frame termination triplet (TAA) lies 20 codons upstream of the ATG. The encoded polypeptide sequence aligns well with other cytochrome P450 sequences and we surmise that the ATG at position 81 represents the correct start site for translation. At the 3' end the truncation of clone 40 lies entirely in the non-coding region downstream of the stop codon. Neither clone contained a poly-A tail but both contain d a potential polyadenylation sequence (AATAAA) at a position corresponding precisely to that seen in the rat cDNA.

2.2 Structure of mouse Hct-1 p lypeptide

As anticipated, nucleotide sequence homology of mouse Hct-1 was highest with human cholesterol 7α-hydroxylase, with approximately 56% identity over the coding region. At the polypeptide level the mouse ORF shows 81% identity to the rat Hct-1 polypeptide over 414 amino acids; the precise degree of similarity may be different as the full protein sequence of rat Hct-1 is not known. Both the human (CYP7) and rat cholesterol 7α-hydroxylase polypeptides share 39% amino acid sequence identity to mouse Hct-1. Fig. 4*A* presents the alignment of mouse Hct-1 polypeptide with human CYP7.

The N-terminus of the Hct-1 polypeptide is hydrophobic, a feature shared by microsomal CYP's. This portion of the polypeptide is thought to insert into the membrane of the endoplasmic reticulum, holding the main bulk of the protein on the cytoplasmic side. Consistent with microsomal CYP's, the N-terminus lacks basic amino acids prior to the hydrophobic core (amino acids 9-34).

Several alignment studies have previously highlighted conserved regions within CYP polypeptides. We therefore inspected the Hct-1 sequence for these conserved regions. CYP's contain a highly conserved motif, FxxGxxxCxG(xxxA), present in 202 of the 205 compiled sequences (Nelson *et al., supra*), that is thought to represent the heme binding site. The arrangement of amino acids around the cysteine residue has been postulated to preserve the three-dimensional structure of this region for ligand binding. This motif is fully conserved in Hct-1 (Fig. 4*B*). A second conserved domain is also present in CYP's responsible for steroid interconversions. While this domain is largely conserved in Hct-1 an invariant Pro residue is replaced, in Hct-1, by Val (Fig. 4*C*); the rat Hct-1 polypeptide also contains a Val residue at this position.

2.3 Expressi n pattern of m use Hct-1

To verify enriched expression of Hct-1 in hippocampus we performed Northern and in situ hybridization analyses on mouse material. In contrast to the

situation in rat, the 1.4 kb clone 12 detected only a 1.8 kb transcript; the 2.1 kb and 5.0 kb transcripts were absent from all tissues examined (Fig. 2*C*). The apparent absence of the 2.1 kb transcript may only reflect a lower abundance of this transcript because at least some mouse cDNA clones extend beyond the upstream polyadenylation site which is thought, in rat, to generate the shorter (1.8 kb) transcript.

To refine this analysis, a 42-mer oligonucleotide was designed according to the DNA sequence of the 3' untranslated region of the cDNA clone upstream of the first polyadenylation site (materials and methods), so as to minimize cross-hybridization with other CYP mRNA's. Coronal sections of mouse brain were hybridized to the ^{35}S-labelled probe and, after emulsion dipping, exposed for autoradiography (Fig. 5). Transcripts were detected throughout mouse brain, with no evidence of restricted expression in the hippocampus (Fig. 5*A,B*). Strongest expression was observed in the corpus callosum, the anterior commisure and fornix while, as in rat, hippocampal expression was particularly prominent in the dentate gyrus (Fig. 5*C*). Moderate expression levels, comparable to those observed in hippocampus, were observed in cerebellum, cortex and olfactory bulb.

2.4 The structure of the mHct-1 gene.

The use of homologous recombination to manipulate the mouse Hct-1 gene requires knowledge of the intron-exon structure of the gene. Sequences upstream of the first Hct-1 exon could also be analysed for elements which contribute to the transcriptional regulation of Hct-1 expression. For these reasons, the organisation of the mouse Hct-1 gene was investigated.

To assess the complexity of the Hct-1 gene in the genome, that is, whether the Hct-1 gene is present as a single copy in the haploid mouse genome, and to assist in mapping of mHct-1 phage clones, the 1.8 kb full length mouse Hct-1 clone was ^{32}P-labelled by random primer labelling and used as a probe on a Southern blot of mouse genomic DNA (Figure 7(a)). Under high stringency conditions the Hct-1 probe recognised a small number of bands within the mouse genomic digests,

suggesting that Hct-1 is present in the mouse genome as a single copy gene. To confirm this, the original 0.3 kb cDNA clone, 14.5a, was used to probe a rat genomic Southern blot. The smaller probe hybridised to a single band in BamHI-, EcoRI-, and XbaI -digested genomic rat DNA (Figure 7(b)).

A mouse genomic DNA library (a gift from A. Reaume, Toronto) prepared from ES cells derived from the 129 mouse strain was screened for genomic clones containing mHct-1 exonic sequence. 750,000 recombinant phage of the *lambda* DASH II library were plated at a density of 50,000 recombinants per 15 cm plate. Duplicate lifts were made and probed with the 1.4 kb rat Hct-1 clone. After the primary screen, 5 clones were isolated. After secondary screening, three of these phage clones were positive and were purified.

Small scale phage DNA was prepared from each phage lysate and cut with NotI to release the inserts. No internal NotI sites were found in any of the clones. Clone I-2 contained a 14 kb insert; clone I-6 contained a 15 kb insert, and clone I-11 contained a 12 kb insert.

These phage clones were mapped by a combination of restriction enzymes which either cut the *lambda* clones rarely, or by using restriction sites found in the mHct-1 cDNA sequence (Figure 3). A 5' probe was created using a 200 bp fragment from the 5' end of mHct-1 cDNA as a probe; this segment extended from the internal BamHI site to an EcoRI site located in the polylinker. The 200 bp 3' cDNA probe extended from the SacI site to the polylinker NotI site. Exon-intron boundaries were determined by subcloning of exon-containing genomic DNA fragments and sequencing (Figure 8).

Phage clones I-6 and I-11 represented 20 kb of contiguous sequence of the Hct-1 locus. I-2 does not overlap withI-6 or I-11, thus the map of the Hct-1 gene in mouse is incomplete. However, the present map shows that mHct-1 spans at least 25 kb of the genome. At least two exons are contained within I-6. The first

exon (referred to as exon II) contains 133 bp of coding sequence, followed by exon III, located 4.0 kb downstream. The 3' boundary of this latter exon is not defined, however approximately 400 bp downstream of its 3' boundary commences exon IV, which together comprise 797 bp of coding sequence. Exon III and IV are also represented in the overlapping sequence of I-11. A fourth exon of at least 345 bp was identified in I-2 (referred to as exon VI). The 3' boundary of this exon has not been identified, thus it is not known whether this contains the remaining coding sequence or if there are additional exons.

The following Table provides a summary of the exon-intron structure of Hct-1 (incomplete) and comparison to human *CYP7* gene structure. * indicates that these exons are not cloned and are not necessarily one exon. ** indicates that the 3' boundary of exon VI is not confirmed and may not necessarily be the final exon.

Exon	cDNA sequence represented	exon size (bp)	*CYP7* exon (bp)
I*	1-142	142	144
II	143-275	133	241
III	276-?	797	587
IV	?-1072	"	131
V*	1073-1246	174	176
VI**	1247-(1821)	(575)	1596

As shown in the Table, cDNA sequence from nucleotides 1073 - 1246 is not represented in the identified exons and must be represented in a separate exon.

142 bp of 5' sequence and 227 bp of 3' sequence have not yet been located in the genomic clones. The remaining 5' sequence is most likely contained in one exon, as the 5' probe (BamHI fragment) consistently recognised two bands by Southern analysis (one of which is exon II sequence). The remaining 3' sequence has not been located and may be part of exon VI or be encoded by a separate exon.

3. ISOLATION OF HUMAN GENOMIC SEQUENCES FOR HCT-1

3.1 Conservation of Hct-1 in humans.

The evolutionary conservation of a gene supports a functionally significant role for that gene in the organism. The conservation of Hct-1 in rodents has been demonstrated by the cloning of the rat and mouse cDNAs for Hct-1. To establish the presence of the Hct-1 gene in the human genome, Southern blotting of human DNA was performed. The rat 1.4 kb clone of Hct-1 was used as a radiolabelled probe and gave strong signals from all three species (Figure 6). A number of hybridising fragments appear to be conserved between species, suggesting conservation of the Hct-1 gene structure. There is a conserved 1.4 kb HindIII band between mouse and rat, while human DNA contains a slightly larger HindIII band of 1.6 kb. Also an EcoRI fragment of 11 kb is conserved in human and rat Hct-1. Conservation of Hct-1 gene structure is also supported from the cDNA digestion patterns of mouse and rat (see Figures 6 and 7), where the SacI, HindIII and PstI sites are conserved between the rodent species.

3.2 A single gene for Hct-1 in mouse, rat and human

Because CYP's comprise a family of related enzymes we wished to determine whether close homologs of Hct-1 are present in the mammalian genome. The rat Hct-1 probe (1.4 kb) was used to probe a genomic Southern blot of rat, mouse and human DNA. In Fig. 6 the probe revealed a simple pattern of cross-hybridizing bands in all DNA's examined. In BamHI-cut human DNA only a single major cross-hybridizing band (4 kb) was detected (Fig. 6), while reprobing with the

300 nt. clone 14-5a yielded, in each lane, a single cross-hybridizing band (not shown). These data argue that a single conserved Hct-1 gene is present in mouse, rat and human, and that the mammalian genome does not contain very close homologs of Hct-1 that would be detected by cross-hybridization (>70-80% homology).

3.3 Isolation of sequences encoding human Hct-1

The rat cDNA clone 14.5a-12 was used to probe a Southern blot of human genomic DNA digested with BamHI according to standard procedures. A single band at 3.8 kb was identified that cross-hybridises with the probe. Accordingly, 20 μg of human genomic DNA was cleaved to completion with BamHI, resolved by agarose gel electrophoresis, and the size range 3.4-4.2 kb selected by reference to markers run on the same gel. The gel fragment was digested by agarase treatment, DNA was purified by phenol extraction and ethanol precipitation, and ligated into BamHI-cut bacteriophage lambda ZAP vector (Stratagene). Following packaging in vitro and plating on a lawn of E. coli strain XL1-Blue , plaque lifts of 100,000 clones were screened for hybridisation to the rat cDNA. 12 positive signals were identified and all contained a 3.8 kb insert. One was selected and the segment was partially sequenced, identifying two regions of high homology to the rat (and mouse) cDNA's and corresponding to exons 3 and 4. Figure 9 presents the nucleotide sequence and Figure 10 compares the human Hct-1 translation product with the cognate mouse polypeptide.

To extend this characterisation, the 3.8 kb BamHI fragment obtained from the size-selected library was used to screen a genomic library of human DNA prepared by partial Sau3A cleavage and insertion of 14-18 kb fragments into a bacteriophage lambda vector according to standard techniques (gift of Dr. P. Estibeiro, CGR). Positive clones were obtained, and restriction mapping of one confirmed that it contains approximately 14 kb of human DNA encompassing the exons identified above and further regions of the Hct-1 gene; together the different genomic clones are thought to encompass the entire Hct-1 gene. The human

genomic sequence may be used to screen human cDNA libraries for full length cDNA clones; alternatively, following complete DNA sequence determination the human genomic sequence may be expressed in mammalian cells by adjoining it to a suitable promoter sequence and cDNA prepared from the correctly spliced mRNA product so produced. Finally, the genomic Hct-1 sequence would permit the entire coding sequence to be deduced so permitting the assembly of a full length Hct-1 coding sequence by de novo synthesis.

3.4 Expression of Hct-1 protein for enzymatic activity analysis

3.4.1. Expression of Hct-1 polypeptide in yeast cells

Recombinant yeast strains are useful vehicles for the production of heterologous cytochrome P450 proteins. It would be possible to express any of the mammalian Hct-1's in yeast, but for simplicity we selected the mouse Hct-1 clone 35. To introduce the mouse Hct-1 (mHct-1) coding sequence into yeast the expression vector pMA91 (Kingsman et al., Meth. Enzymol. 185: 329-341, 1990) was employed. The unique BglII site in pMA91 was converted to a NotI site by inserting the oligonucleotide 5'GATCGCGGCCGC3' according to standard procedures. Following cleavage of the resulting plasmid (pMA91-Not) with NotI the mHct-1 cDNA clone 35 was introduced, placing mHct-1 expression under the control of the yeast PGK (phosphoglycerokinase) promoter for high level expression in yeast cells (Figure 12A). A similar construct utilising the mHct-1 cDNA clone 35 is depicted in Figure 12B. Expression of mHct-1 in yeast using these plasmid permits the purification of the protein and determination of substrate specificity.

3.4.2. Expression of Hct-1 polypeptide in vaccinia virus

Expression in vaccinia virus is a routine procedure and has been widely employed for the expression of heterologous cytochromes P450 in mammalian cells, including HepG2 and Hela cells (Gonzalez, Aoyama and Gelboin, Meth. in Enzymol. 206: 85-92, 1991; Waxman et al., Archives Biochem. Biophys 290, 160-166,1991). Accordingly we selected plasmid pTG186-poly (Lathe et al.,

Nature 326, 878-880, 1987) as the transfer/expression vector, although other similar vectors are widely available and may also be employed.

To demonstrate the expression of mammalian Hct-1's in vaccinia virus, for simplicity we selected the mHct-1 clone 35. Similar techniques are applicable to rat and human Hct-1's. To enhance expression we elected to modify the 5' end to conform better to the translation consensus for mammalian cells (YYAYYATGR) though this modification may not be essential.

Accordingly, two oligonucleotides were designed corresponding to the 5' and 3' regions of the mouse cDNA.

The 5' oligonucleotide:

(5'-GGCCCTCGAGCCACCATGCAGGGGAGCCACG-3')

is homologous to the region surrounding the translation initiation site but converts the sequence immediately prior to the ATG to the sequence CCACC; in addition, the oligonucleotide contains a XhoI restriction site for subsequent cloning. The 3' oligonucleotide (GGCCGAATTCTCAGCTTCTCCAAGAA) was chosen according to the sequence downstream of the translation stop site and contains, in addition, an EcoRI site for subsequent cloning. These oligonucleotides were employed in polymerase chain reaction (PCR) amplification through 5 cycles on the clone 35 template; the products were applied to an agarose gel and the desired product band at 1.65 kb was cut out and extracted by standard procedures.

Following cleavage with XhoI and EcoRI the modified fragment was introduced between the EcoRI and SalI sites of pTG186-poly, generating pVV-mHct-1. Recombinational exchange was used to transfer the expression vector to the vaccinia virus genome according to standard procedures, generating VV-mHct-1, as depicted in Figure 13. This recombinant will permit the expression

of high levels of mHct-1 and the identification of the substrate specificity of the protein, as well as the production of antibodies directed against mHct-1.

To identify the product of P450-mediated metabolism, microsomes may easily be prepared (Waxman, Biochem. J. 260: 81-85, 1989) from vaccinia-infected cells: these are incubated with labelled precursors, eg. steroids, and the product identified by thin layer chromatography according to standard procedures (Waxman, Methods in Enzymology 206:462-476).

The Hct-1 provided according to this invention thereby provides a route for the large-scale production of the product described above, for instance a modified steroid, by expressing the P450 in a recombinant organism and supplying the substrate for conversion. It will also be possible to engineer recombinant yeast, for instance, to synthesise the substrate for the Hct-1 P450 in vivo, so as to allow production of the Hct-1 product from yeast supplied with a precursor, for instance cholesterol or other molecule, if that yeast is engineered to contain other P450's or modifying enzymes. It may be possible for Hct-1 to act on endogenous sterols and steroids in yeast to yield product.

Finally, the Hct-1 product may be part of a metabolic chain, and recombinant organisms may be engineered to contain P450's or other enzymes that convert the Hct-1 product to a subsequent product that may in turn be harvested from the organism.

4. DISCUSSION

In experiments to characterize transcripts enriched in the hippocampal formation we isolated cDNA clones corresponding to Hct-1 (hippocampal transcript) from a library prepared from rat hippocampus RNA. In rat, expression appeared to be most abundant in hippocampus with some expression in cortex and substantially less expression other in brain regions. Elsewhere in the body transcripts were only

detected in liver and, to a lesser extent, in kidney; expression was barely detectable in ovary, testis and adrenal, also sites of steroid transformations. Hepatic expression was sexually dimorphic with Hct-1 mRNA barely detectable in female liver. In rat brain and liver, Hct-1 identifies two transcripts of 1.8 and 2.1 kb that appear to be generated by alternative polyadenylation; a 5.0 kb transcript weakly detected in brain is thought not to originate from the Hct-1 gene but instead encodes a polypeptide related to the GTPase activating protein, ABR (active BCR-related).

Sequence analysis of Hct-1 cDNA clones revealed an extensive open reading frame encoding a protein with homology to cytochromes P450 (CYP's), a family of heme-containing mono-oxygenases responsible for a variety of steroid and fatty acid interconversions and the oxidative metabolism of xenobiotics. Although the mouse cDNA coding region appears complete, the absence of a consensus translation initiation site flanking the presumed initiation codon could indicate that Hct-1 polypeptide synthesis is subject to regulation at the level of translation initiation.

Homology was highest with rat and human cholesterol 7α-hydroxylase, known as CYP7. While related, Hct-1 is clearly distinct from CYP7, sharing only 39% homology over the full length of the protein. CYP polypeptides sharing greater than 40% sequence identity are generally regarded as belonging to the to the same family, and Hct-1 and CYP7 (39% similarity) are hence borderline. The conservation of other unique features between Hct-1 and CYP7 however argues for a close relationship and Hct-1 has been provisionally named 'CYP7B' by the P450 Nomenclature Committee (D.R. Nelson, personal communication).

From the Hct-1 leader sequence we surmise that the Hct-1 polypeptide resides, like CYP7, in the endoplasmic reticulum and not in mitochondria, the other principal cellular site of CYP activity. The strictly conserved heme binding site motif FxxGxxxCxG(xxxA) is clearly present in Hct-1 (residues 440-453). It is of note that the 'steroidogenic domain', conserved in many CYP's responsible for

steroid interconversions, is also present in Hct-1 (amino acids 348-362), except that a consensus Pro residue is replaced by Val in both the mouse and rat Hct-1 polypeptides. Of previously known 34 CYP sequences, only 4 contain an amino acid residue other than Pro at this position. Whereas 2 of these harbour an unrelated amino acid (Glu; CYP3A1, CYP3A3), interestingly, a Val residue is present in bovine CYP17 (steroid 17α-hydroxylase, 44) at a position equivalent to that in Hct-1 while human CYP17 harbours a conservative substitution at this site (Leu; 44). Despite this similarity, however, the overall extent of homology between Hct-1 and CYP17 (22.5%, not shown) is lower than with CYP7 (39%)

Neither Hct-1 and CYP7 appear to contain a conserved O_2 binding pocket (equivalent to residues 285-301 in Hct-1). Crystallographic studies on the bacterial CYP101 indicated that a Thr residue (corresponding to position 294 in Hct-1) disrupts helix formation in that region and is important in providing a structural pocket for an oxygen molecule. Site-directed mutagenesis of this Thr residue in both CYP4A1 and CYP2C11 demonstrated that this region can influence substrate specificity and affinity. In both Hct-1 and CYP7 the conserved Thr residue is replaced by Asn. This modification suggests that Hct-1 and CYP7 are both structurally distinct from other CYP's in this region; this may be reflected both in modified oxygen interaction and substrate choice.

The sexual dimorphism of Hct-1 expression observed in rat resembles that observed with a number of other CYP's. CYP2C12 is expressed preferentially in liver of the female rat while, like Hct-1, CYP2C11 is highly expressed in male liver but only at low levels in the female tissue. This dimorphic expression pattern of CYP2C family members is thought to be determined by the dimorphism of pulsatility of growth hormone secretion. Brain expression of Hct-1 is not subject to this control suggesting that regulatory elements determining Hct-1 expression in brain differ from those utilized in liver. However, we have not examined species other than rat; it cannot be assumed that the same regulation will exist in other species. Indeed, sexually dimorphic gene expression is not necessarily conserved between different strains of mouse.

Expression of Hct-1 was widespread in mouse brain. The expression pattern was most consistent with glial expression but further experiments will be required to compare neuronal and non-neuronal levels of expression. In mouse brain only the 1.8 kb transcript was detected, though cDNA's were obtained corresponding to transcripts extending beyond the first polyadenylation site; such extended transcripts are thought to give rise to the 2.1 kb transcript in rat. This suggests the downstream polyadenylation site seen in rat Hct-1 is under-utilized in mouse Hct-1 or absent. While in situ hybridization studies of Hct-1 in rat brain were inconclusive, a difference in expression pattern between mouse and rat appears likely; further work will be required to confirm this. However, such a difference would be unsurprising because cytochromes P450 are well known to vary widely in their level and pattern of expression in different species; for instance, hepatic testosterone 16-hydroxylation levels differ by more than 100-fold between guinea pig and rat.

Our data indicate that the Hct-1 gene is present in rat, mouse and human, and there appear to be no very close relatives in the mammalian genome. While CYP genes are scattered over the mouse and human genomes, CYP subfamilies can cluster on the same chromosome. For instance, the human CYP2A and 2B subfamily genes are linked to chromosome 19, CYP2C and 2E subfamilies are located on human chromosome 10, and the mouse cyp2a, 2b and 2e subfamilies are present on mouse chromosome 7. The gene encoding human cholesterol 7α-hydroxylase (CYP7) is located on chromosome 8q11-q12.

Together our data argue that Hct-1 and CYP7 are closely related: this suggests that the substrate for Hct-1, so far unknown, is likely to be related to cholesterol or one of its steroid metabolites. This interpretation is borne out by the presence, in Hct-1, of the steriodogenic domain conserved in a number of steroid-metabolizing CYP's. While experiments are underway to determine the substrate specificity of Hct-1, the possibility that Hct-1 acts on cholesterol or its steroid metabolites in brain is of some interest. CYP7 (cholesterol 7α-hydroxylase) is

responsible for the first step in the metabolic degradation of cholesterol. This is of note in view of the association of particular alleles of the APOE gene encoding the cholesterol transporter protein apolipoprotein E with the onset of Alzheimer's disease, a neurodegenerative condition whose cognitive impairments are associated with early dysfunction of the hippocampus.

What role might Hct-1 play in the brain? In the adult CYP's are generally expressed abundantly in liver, adrenal and gonads, while the level of CYP activity in brain is estimated to be 0.3 to 3% of that found in liver (see 58). Because levels of Hct-1 mRNA expression in rat and mouse brain far exceed those in liver it could be argued that the primary function of Hct-1 lies in the central nervous system. The documented ability of cholesterol-derived steroids to interact with neurotransmitter receptors and modulate both synaptic plasticity and cognitive function suggests that Hct-1 and its metabolic product(s) may regulate neuronal function *in vivo*.

5. SUMMARY

Hct-1 (hippocampal transcript) was detected in a differential screen of a rat hippocampal cDNA library. Expression of Hct-1 was enriched in the formation but was also detected in rat liver and kidney, though at much lower levels; expression was barely detectable in testis, ovary and adrenal. In liver, unlike brain, expression was sexually dimorphic: hepatic expression was greatly reduced in female rats. In mouse, brain expression in was widespread, with the highest levels being detected in corpus callosum; only low levels were detected in liver. Sequence analysis of rat and mouse Hct-1 cDNAs revealed extensive homologies with cytochrome P450's (CYP's), a diverse family of heme-binding monooxygenases that metabolize a range of substrates including steroids, fatty acids and xenobiotics. Among the CYP's, Hct-1 is most similar (39% at the amino acid sequence) to cholesterol 7α-hydroxylase (CYP7), and contains the diagnostic steriodogenic domain present in other steriod-metabolizing CYPs, but clearly represents a type of CYP not previously reported. Genomic Southern analysis indicates that a single gene

corresponding to Hct-1 is present in mouse, rat and human. Hct-1 is unusual in that, unlike all other CYP's described, the primary site of expression is in the brain. Similarity to CYP7 and other steroid-metabolizing CYP's argues that Hct-1 plays a role in steroid metabolism in brain, notable because of the documented ability of brain-derived steroids (neurosteroids) to modulate cognitive function in vivo.

6. DETAILS OF EXPERIMENTAL PROTOCOLS

Northern analysis--Total RNA was extracted by tissue homogenization in guanidinium thiocyanate according to a standard procedure and further purified by centrifugation through a CsCl cushion. Where appropriate, polyA-plus RNA was selected on oligo-dT cellulose. Electrophoresis of RNA (10 μg) on 1% agarose in the presence of 7% formaldehyde was followed by capillary transfer to nylon membranes, baking (2 h, 80°C), and rinsing in hybridization buffer (0.25 M NaPhosphate, pH 7.2; 1 mM EDTA, 7% sodium dodecyl sulphate [SDS], 1% bovine serum albumin) as described (Church *et al., supra*). Probes were prepared by random-priming of DNA polymerase copying of denatured double-stranded DNA. Hybridization (16 h, 68°C) was followed by washing (3 times, 20 mM NaPhosphate pH 7.2, 1 mM EDTA, 1% SDS, 20 min.) and membranes exposed for autoradiography. The loading control probe was a 0.5 kb cDNA encoding the ubiquitously expressed rat ribosomal protein S26.

In situ hybridization--Synthetic Hct-1 oligonucleotide probes 5'-dGACAGGTTTTGTGACCCAAAACAAACTGGATGGATCGCAATC-3' (rat, 55% G+C) and 5'-ATCACGGAGCTCAGCACATGCAGCCTTACTCTGCAAAGCTTC-3' (mouse - 48% G+C) were labelled using terminal transferase (Boehringer Mannheim) and α-^{35}S-dATP (Amersham) according to the manufacturer's instructions.

The control probe,

5'-dAGCCTTCTGGGTCGTAGCTGACTCCTGCTGCTGAGCTGCAACAGCTTT-3'

(56% G+C) was based on human opsin cDNA. Frozen coronal 10 μm sections of brain were fixed (4% paraformaldehyde, 10 min), rinsed, treated with proteinase K (20 μg/ml in 50 mM Tris.HCl, pH 7.4, 5 mM EDTA, 5 min), rinsed, and refixed with paraformaldehyde as before. Following acetylation (0.25% acetic anhydride, 10 min) and rinsing, sections were dehydrated by passing though increasing ethanol concentrations (30, 50, 70, 85, 95, 100, 100%, each for 1 minute except the 70% step [5 min]). Following $CHCl_3$ treatment (5 min), and rinsing in ethanol, sections were dried before hybridization. Hybridization in buffer (4 x standard saline citrate [1 x SSC = 0.15 M NaCl, 0.015 M Na_3citrate], 50% v/v formamide, 10% w/v dextran sulphate, 1x Denhardt's solution, 0.1% SDS, 500 μg/ml denatured salmon sperm DNA, 250 μg/ml yeast tRNA) was for 16 h at 37°C. Slides were washed (4 x 15 min., 1xSSC, 60°C; 2 x 30 min., 1 x SSC, 20°C), dipped into photographic liquid emulsion (LM-1, Amersham), exposed and developed according to the manufacturer's specifications. Slides were counterstained with 1% methyl green.

Southern hybridization--Genomic DNA prepared from mouse or rat liver, or from human lymphocytes, was digested with the appropriate restriction endonuclease, resolved by agarose gel electrophoresis (0.7%) and transferred to Hybond-N membranes. Following baking (2 h, 80°C), hybridization conditions were as described for Northern analysis.

Hybridisation Conditions. Hybridisation conditions used were based on those described by Church and Gilbert, Proc. Natl. Acad. Sci. USA (1984) 81, 1991-1995.

1. Filters were pre-wet in 2XSSC.
2. The hybridisation was performed in a rotating glass cylinder (Techne Hybridiser ovens). 10 ml of Hybridisation Buffer was added to the cylinder with the filter.

3. Prehybridisation and hybridisation were carried out at 68°C unless otherwise specified.

4. The filters were prehybridised for 30 minutes, after which the probe was added directly and hybridisation proceeded overnight. (Double-stranded probes were denatured by boiling for 2 minutes, then placing on ice).

5. Washes were performed at 68°C (unless otherwise stated) with 2 changes of Wash Buffer I for 10 minutes each, followed by three changes of Wash Buffer II each for 20 minutes.

6. The filters were blotted dry, but not allowed to dry out, then placed between Saran wrap, and against X-ray film for autoradiography.

Hybridisation Buffer:

0.25 M sodium phosphate pH 7.2
1 mM EDTA
7% SDS
1% BSA

Wash Buffer I:

20 mM sodium phosphate pH 7.2
2.5% SDS
0.25% BSA
1 mM EDTA

Wash Buffer II:

20 mM sodium phosphate pH 7.2
1 mM EDTA
1% SDS

Screening of Bacteriophage lambda libraries. The rat hippocampus cDNA library was oligo-(dT)-NotI primed and cloned in *lambda* ZAP II (Stratagene) with an EcoRI adaptor at the 5' end, and was prepared in the lab by Miss M. Richardson and Dr. J. Mason; the mouse liver cDNA library was oligo-(dT)-primed and cloned into *lambda* gt10 with EcoRI/NotI adaptors, and was a gift from Dr. B. Luckow, Heidelberg; the mouse ES cell genomic library was cloned from a partial Sau3A digest into *lambda* DASH II (Stratagene), and was a gift from A. Reaume, Toronto.

The libraries were screened as described above by hybridization.

In vivo excision of pBluescript from *lambda* ZAP II vector was performed using the ExAssist/SOLR system (Stratagene, 200253).

In situ hybridisation. Frozen 10μ coronal sections of rat and mouse brains were provided by Dr. M. Steel.

Hybridisation Conditions All probes were oligonucleotides which were labelled by homopolymer tailing using a-^{35}S-dATP and terminal transferase.

The sequences or references of the oligonucleotides used as probes for *in situ* hybridisation were as follows:

rat Hct-1 (a 45-mer, beginning 26 nt 5' from the polyA tail, nucleotides 1361-1403 in Figure 4.2) (for relative position in mouse gene, see Figure 4.3)
5'-GACAGGTTTTGTGACCCAAAACAAACTGGATGGATCGCAATC-3'

Nathans mouse Hct-1 (nt 1558-1599)
5'-ATCACGGAGCTCAGCACATGCAGCCTTACTCTGCAAAGCTTC-3'

rat clone 13 (a 42-mer, beginning 112 nt 5' from polyA tail)
5'-TATATCCATACCAACTTATTGGGAGTCCCATCCTACCTCATCAGC-3'

rat/mouse muscarinic receptor M1 (Buckley *et al.*, 1988) .

rat/mouse opsins (Nathans et al., Science (1986) **232**, 193-202)

1. The prepared ^{35}S -tailed probe (resuspended in 10 mM DTT in TE) was diluted to 2 x 10^6 cpm/ml in hybridisation buffer. DTT is also added to this mixture to a final concentration of 50 mM.
2. 100 ml of the probe mixture was carefully layered onto each microscope slide. A piece of parafilm cut to the size of the microscope slide was then layered over the probe mixture, allowing the probe and hybridisation mixture to cover all the sections. Air bubbles under the parafilm were avoided.
3. The slides were placed in a humidified container, sealed, and incubated at 37°C overnight.
4. After hybridisation, the parafilm was carefully removed using forceps.
5. The slides were placed back in Coplin jars, and the hybridised sections washed in four changes of 1XSSC for 15 minutes at 55°C or 60°C, and then two changes of 1XSSC for 30 minutes at room temperature.
6. The slides were rinsed briefly in dH_2O, then left to air dry.

Hybridisation Buffer:*

4XSSC
50% (v/v) deionised formamide
10% (w/v) dextran sulphate
1X Denhardt's solution
0.1% (w/v) SDS
500 µg/ml ssDNA
250 µg/ml yeast tRNA
*buffer was de-gassed before use

7. FIGURE LEGENDS

FIG.1. **Sequence of partial rat Hct-1 cDNA and the encoded polypeptide.** The nucleotide sequence and translation product of the 1.4 kb cDNA clone 12 including additional clone 7 sequence (lower case). The two putative polyadenylation signals are underlined.

FIG. 2. **Northern analysis of Hct-1 expression in adult rat and mouse brain.** *Panel A.* Expression in rat brain and other tissues; *panel B.* sexually dimorphic expression in rat liver; *panel C.* Expression in mouse tissues. Poly-A^+ (A) or total (B,C) RNA from organs of adult animals were resolved by gel electrophoresis; the hybridization probe was rat Hct-1 cDNA clone 12 (1.4 kb), the probe for the loading control (below) corresponds to ribosomal protein S26. Tissues analysed are: Hi, hippocampus; RB, remainder of brain lacking hippocampus; Cx, cortex; Cb, cerebellum; Ob; olfactory bulb; Li, liver; He, heart; Th, thymus; Ki, kidney; Ov, ovary; Te, testis; Lu, lung.

FIG. 3. **Mouse Hct-1 cDNA and the sequence of the encoded polypeptide.** The restriction map of the cDNA (above) corresponds to the compilation of two independent clones sequenced; the cross-hatched box indicates the coding region. The nucleotide sequence and translation product (below) derives from this compilation. Lower case sequences indicate the 59 additional 5′ nucleotides in clone 40 and the 99 additional 3′ nucleotides in clone 35. The putative polyadenylation site is underlined.

FIG. 4. **Alignment of mouse Hct-1 with human CYP7 (ch lesterol 7*α*-hydr xylase, Noshiro and Okuda, 1990) and other steroidogenic P450s.**
Panel A: Identical amino acids are indicated by a bar; hyphens in the amino acid sequences indicate gaps introduced during alignment. The N-terminal hydrophobic leader sequences are underlined. The position of the conserved Thr residue within the O_2-binding pocket of other CYP's (43), but replaced by Asn in Hct-1 (position 294) and CYP7, is indicated by an asterisk. *Panels B,C:* conserved residues in the heme-binding (residues 440-453, *B*) and steroidogenic (residues 348-362, *C*) domains conserved between Hct-1 and other similar CYP's (overlined in *A*). Sequences are human CYP7 (7*α*-hydroxylase; 37); bovine CYP17 (17*α*-hydroxylase; 44); human CYP11B1 (steroid ß-hydroxylase; 45); human CYP21B (21-hydroxylase; 11); human CYP11A1 (P450scc; cholesterol side-chain cleavage; 46); human CYP27 (27-hydroxylase; 47).

FIG. 5. **Analysis of Hct-1 expression in adult mouse brain.**
The hybridization probe was a synthetic oligonucleotide corresponding to the 3' untranslated region of mouse Hct-1 cDNA. *Panel a:* coronal section; *panel b:* coronal section, rostral to *a*, showing hybridization in corpus callosum, cc; fornix, f; and anterior commissure, ac; *panel c:* enlargement of section through the hippocampus; DG, dentate gyrus; *panel d:* section adjacent to the section in *a* hybridized with an oligonucleotide specific for opsin (negative control).

FIG. 6. **Southern analysis of Hct-1 coding sequences in mouse, rat and human**
Total DNA was cleaved as indicated with restriction endonucleases B, BamHI; E, EcoRI; H, HindIII; X, XbaI; resolved by agarose gel electrophoresis, and probed with rat Hct-1 cDNA clone 12 before exposure to autoradiography.

FIG. 7 **Genomic DNA Southern blot analysis of Hct-1** (a) Mouse genomic DNA probed with the full-length mouse Hct-1 cDNA clone. (b) Rat genomic DNA probed with clone 14.5a (original 0.3 kb clone of rHct-1). 10 μg of genomic DNA was digested with the indicated enzymes.

FIG. 8 **Gen mic map of mouse Hct-1 (incomplete)**. Exons II, III, IV and VI are represented on the phage clones (filled boxes). Exons I and V are not located. As indicated in Table 4.1, the boundaries of exons II, III B (BamHI); H(HindIII); S(SacI); X(XhoI)

SEQUENCE ID NO: 1

```
A  L  E  Y  Q  Y  V  M  K  N  P  K  Q  L  S  F  E  K  F  S
GCCTTGGAGTACCAGTATGTAATGAAAAACCCAAAACAATTAAGCTTTGAGAAGTTCAGC   60
R  R  L  S  A  K  A  F  S  V  K  K  L  L  T  N  D  D  L  S
CGAAGATTATCAGCGAAAGCCTTCTCTGTCAAGAAGCTGCTAACTAATGACGACCTTAGC  120
N  D  I  H  R  G  Y  L  L  L  Q  G  K  S  L  D  G  L  L  E
AATGACATTCACAGAGGCTATCTTCTTTTACAAGGCAAATCTCTGGATGGTCTTCTGGAA  180
T  M  I  Q  E  V  K  E  I  F  E  S  R  L  L  K  L  T  D  W
ACCATGATCCAAGAAGTAAAAGAAATATTTGAGTCCAGACTGCTAAAACTCACAGATTGG  240
N  T  A  R  V  F  D  F  C  S  S  L  V  F  E  I  T  F  T  T
AATACAGCAAGAGTATTTGATTTCTGTAGTTCACTGGTATTTGAAATCACATTTACAACT  300
I  Y  G  K  I  L  A  A  N  K  K  Q  I  I  S  E  L  R  D  D
ATATATGGAAAAATTCTTGCTGCTAACAAAAAACAAATTATCAGTGAGCTGAGGGATGAT  360
F  L  K  F  D  D  H  F  P  Y  L  V  S  D  I  P  I  Q  L  L
TTTTTAAAATTTGATGACCATTTCCCATACTTAGTATCTGACATACCTATTCAGCTTCTA  420
R  N  A  E  F  M  Q  K  K  I  I  K  C  L  T  P  E  K  V  A
AGAAATGCAGAATTTATGCAGAAGAAAATTATAAAATGTCTCACACCAGAAAAAGTAGCT  480
Q  M  Q  R  R  S  E  I  V  Q  E  R  Q  E  M  L  K  K  Y  Y
CAGATGCAAAGACGGTCAGAAATTGTTCAGGAGAGGCAGGAGATGCTGAAAAAATACTAC  560
G  H  E  E  F  E  I  G  A  H  H  L  G  L  L  W  A  S  L  A
GGGCATGAAGAGTTTGAAATAGGAGCACATCATCTTGGCTTGCTCTGGGCCTCTCTAGCA  600
N  T  I  P  A  M  F  W  A  M  Y  Y  L  L  Q  H  P  E  A  M
AACACCATTCCAGCTATGTTCTGGGCAATGTATTATCTTCTTCAGCATCCAGAAGCTATG  660
E  V  L  R  D  E  I  D  S  F  L  Q  S  T  G  Q  K  K  G  P
GAAGTCCTGCGTGACGAAATTGACAGCTTCCTGCAGTCAACAGGTCAAAAGAAAGGACCT  720
G  I  S  V  H  F  T  R  E  Q  L  D  S  L  V  C  L  E  S  A
GGAATTTCTGTCCACTTCACCAGAGAACAATTGGACAGCTTGGTCTGCCTGGAAAGCGCT  780
I  L  E  V  L  R  L  C  S  Y  S  S  I  I  R  E  V  Q  E  D
ATTCTTGAGGTTCTGAGGTTGTGCTCCTACTCCAGCATCATCCGTGAAGTGCAAGAGGAT  840
M  D  F  S  S  E  S  R  S  Y  R  L  R  K  G  D  F  V  A  V
ATGGATTTCAGCTCAGAGAGTAGGAGCTACCGTCTGCGGAAAGGAGACTTTGTAGCTGTC  900
F  P  P  M  I  H  N  D  P  E  V  F  D  A  P  K  D  F  R  F
TTTCCTCCAATGATACACAATGACCCAGAAGTCTTCGATGCTCCAAAGGACTTTAGGTTT  960
D  R  F  V  E  D  G  K  K  K  T  T  F  F  K  G  G  K  K  L
GATCGCTTCGTAGAAGATGGTAAGAAGAAAACAACGTTTTTCAAAGGAGGAAAAAAGCTG 1020
K  S  Y  I  I  P  F  G  L  G  T  S  K  C  P  G  R  Y  F  A
AAGAGTTACATTATACCATTTGGACTTGGAACAAGCAAATGTCCAGGCAGATACTTTGCA 1080
I  N  E  M  K  L  L  V  I  I  L  L  T  Y  F  D  L  E  V  I
ATTAATGAAATGAAGCTACTAGTGATTATACTTTTAACTTATTTTGATTTAGAAGTCATT 1140
```

SequID No: 1 (contd)

```
D  T  K  P  I  G  L  N  H  S  R  M  F  L  G  I  Q  H  P  D
GACACTAAGCCTATAGGACTAAACCACAGTCGCATGTTTCTGGGCATTCAGCATCCAGAC 1200
S  D  I  S  F  R  Y  K  A  K  S  W  R  S  ***
TCTGACATCTCATTTAGGTACAAGGCAAAATCTTGGAGATCCTGAAAGGGTGGCAGAGAA 1260

GCTTAGCGGAATAAGGCTGCACATGCTGAGCTCTGTGATTTGCTGTACTCCCCAAATGCA 1320

GCCACTATTCTTGTTTGTTAGAAAATGGCAAATTTTTATTTGATTGCGATCCATCCAGTT 1380

TGTTTTGGGTCACAAAACCTGTCATAAAATAAAGCGCTGTCATGGTGTaaaaaaatgtca 1440

tggcaatcatttcaggataaggtaaaataacgttttcaagtttgtacttactatgatttt 1500

tatcatttgtagtgaatgtgcttttccagtaataaatttgcgccagggtgatttttttta 1560

attactgaaatcctctaatatcggttttatgtgctgccagaaaagtgtgccatcaatgga 1620

cagtataacaatttccagttttccagagaagggagaaattaagccccatgagttacgctg 1680

tataaaattgttctcttcaactataatatcaataatgtctatatcaccaggttacctttg 1740

cattaaatcgagttttgcaaaag 1763
```

SequID No: 2

```
ggcaggcacagcctctggtctaagaagagagggcactgtgcagaagccatcgctccctaC    60
                      M  Q  G  A  T  T  L  D  A  A  S  P  G  P    14
AGAGCCGCCAGCTCGTCGGGATGCAGGGAGCCACGACCCTAGATGCCGCCTCGCCAGGGC   120
  L  A  L  L  G  L  L  F  A  A  T  L  L  L  S  A  L  F  L  L    34
CTCTCGCCCTCCTAGGCCTTCTCTTTGCCGCCACCTTACTGCTCTCGGCCCTGTTCCTCC   180
  T  R  R  T  R  R  P  R  E  P  P  L  I  K  G  W  L  P  Y  L    54
TCACCCGGCGCACCAGGCGCCCTCGTGAACCACCCTTGATAAAAGGTTGGCTTCCTTATC   240
  G  M  A  L  K  F  F  K  D  P  L  T  F  L  K  T  L  Q  R  Q    74
TTGGCATGGCCCTGAAATTCTTTAAGGATCCGTTAACTTTCTTGAAAACTCTTCAAAGGC   300
  H  G  D  T  F  T  V  F  L  V  G  K  Y  I  T  F  V  L  N  P    94
AACATGGTGACACTTTCACTGTCTTCCTTGTGGGGAAGTATATAACATTTGTTCTGAACC   360
  F  Q  Y  Q  Y  V  T  K  N  P  K  Q  L  S  F  Q  K  F  S  S   114
CTTTCCAGTACCAGTATGTAACGAAAAACCCAAAACAATTAAGCTTTCAGAAGTTCAGCA   420
  R  L  S  A  K  A  F  S  V  K  K  L  L  T  D  D  D  L  N  E   134
GCCGATTATCAGCGAAAGCCTTCTCTGTAAAGAAGCTGCTTACTGATGACGACCTTAATG   480
  D  V  H  R  A  Y  L  L  L  Q  G  K  P  L  D  A  L  L  E  T   154
AAGACGTTCACAGAGCCTATCTACTTCTACAAGGCAAACCTTTGGATGCTCTTCTGGAAA   540
  M  I  Q  E  V  K  E  L  F  E  S  Q  L  L  K  I  T  D  W  N   174
CTATGATCCAAGAAGTAAAAGAATTATTTGAGTCCCAACTGCTAAAAATCACAGATTGGA   600
  T  E  R  I  F  A  F  C  G  S  L  V  F  E  I  T  F  A  T  L   194
ACACAGAAAGAATATTTGCATTCTGTGGCTCACTGGTATTTGAGATCACATTTGCGACTC   660
  Y  G  K  I  L  A  G  N  K  K  Q  I  I  S  E  L  R  D  D  F   214
TATATGGAAAAATTCTTGCTGGTAACAAGAAACAAATTATCAGTGAGCTAAGGGATGATT   720
  F  K  F  D  D  M  F  P  Y  L  V  S  D  I  P  I  Q  L  L  R   234
TTTTTAAATTTGATGACATGTTCCCATACTTAGTATCTGACATACCTATTCAGCTTCTAA   780
  N  E  E  S  M  Q  K  K  I  I  K  C  L  T  S  E  K  V  A  Q   254
GAAATGAAGAATCTATGCAGAAGAAAATTATAAAATGCCTCACATCAGAAAAAGTAGCTC   840
  M  Q  G  Q  S  K  I  V  Q  E  S  Q  D  L  L  K  R  Y  Y  R   274
AGATGCAAGGACAGTCAAAAATTGTTCAGGAAAGCCAAGATCTGCTGAAAAGATACTATA   900
  H  D  D  P  E  I  G  A  H  H  L  G  F  L  W  A  S  L  A  N   294
GGCATGACGATTCTGAAATAGGAGCACATCATCTTGGCTTTCTCTGGGCCTCTCTAGCAA   960
  T  I  P  A  M  F  W  A  M  Y  Y  I  L  R  H  P  E  A  M  E   314
ACACCATTCCAGCTATGTTCTGGGCAATGTATTATATTCTTCGGCATCCTGAAGCTATGG  1020
  A  L  R  D  E  I  D  S  F  L  Q  S  T  G  Q  K  K  G  P  G   334
AAGCCCTGCGTGACGAAATTGACAGTTTCCTGCAGTCAACAGGTCAAAAGAAAGGGCCTG  1080
  I  S  V  H  F  T  R  E  Q  L  D  S  L  V  C  L  E  S  T  I   354
GAATTTCAGTCCACTTCACCAGAGAACAATTGGACAGCTTGGTCTGCCTGGAAAGCACTA  1140
```

SequID No: 2 (contd)

```
  L  E  V  L  R  L  C  S  Y  S  S  I  I  R  E  V  Q  E  D  M  374
TTCTTGAGGTTCTGAGGCTGTGCTCATACTCCAGCATCATCCGAGAAGTGCAGGAGGATA 1200
  N  L  S  L  E  S  K  S  F  S  L  R  K  G  D  F  V  A  L  F  394
TGAATCTCAGCTTAGAGAGTAAGAGTTTCTCTCTGCGGAAAGGAGATTTTGTAGCCCTCT 1260
  P  P  L  I  H  N  D  P  E  I  F  D  A  P  K  E  F  R  F  D  414
TTCCTCCACTCATACACAATGACCCGGAAATCTTCGATGCTCCAAAGGAATTTAGGTTCG 1320
  R  F  I  E  D  G  K  K  K  S  T  F  F  K  G  G  K  R  L  K  434
ATCGGTTCATAGAAGATGGTAAGAAGAAAAGCACGTTTTTCAAAGGAGGGAAGAGGCTGA 1380
  T  Y  V  M  P  F  G  L  G  T  S  K  C  P  G  R  Y  F  A  V  454
AGACTTACGTTATGCCTTTTGGACTCGGAACAAGCAAATGTCCAGGGAGATATTTTGCAG 1440
  N  E  M  K  L  L  L  I  E  L  L  T  Y  F  D  L  E  I  I  D  474
TGAACGAAATGAAGCTACTGCTGATTGAGCTTTTAACTTATTTTGATTTAGAAATTATCG 1500
  R  K  P  I  G  L  N  H  S  R  M  F  L  G  I  Q  H  P  D  S  494
ACAGGAAGCCTATAGGGCTAAATCACAGTCGGATGTTTTTAGGTATTCAGCACCCCGATT 1560
  A  V  S  F  R  Y  K  A  K  S  W  R  S  ***                  507
CTGCCGTCTCCTTTAGGTACAAAGCAAAATCTTGGAGAAGCTGAAAGTGTGGCAGAGAAG 1620

CTTTGCAGAGTAAGGCTGCATGTGCTGAGCTCCGTGATTTGGTGCACTCCCCCAAATGCA 1680

ACCGCTACTCTTGTTTGAAAATGGCAAATTTATATTTGGTTGAGATCAATCCAGTTGGTT 1740

TTGGGTCACAAAACCTGTCATAAAATAAAGCAGTGTGATGGtttaaaaaatgtcatggca 1800

atcatttcaggataaggtaaaataacattttcaagtttgtacttactatgatttttatca 1860

                                                              1880
tttgtagtgaatgtgctttt
```

SequID No: 3.

```
ggatccaaccaagtttccagatcttataaatgtggtgaatggtgaatgacttcctgaaga   60

atggatgaatggatgtgttctagtttggaatcctgtgtcagtcacaagtcaatatgtgac  120

cttgaacatgttattaaatctcccacatccataaaagtgaaaatgctggcattagtggat  180

ttttgccagtgttgaattagacatttatttgtgagtacctgctccatacagtatggtcat  240

ttatttgagttaaaattgttgtatttgaacaaaactcagatgacacctaagcatgaaaaa  300
                                                    intron 2
gctctttatgaagtataaatactcagaaatggaatggcatgttgccaatttgttttctgc  360

tttattgagggaaatatatgagaagtatttaagtcaggggattatgaggaatatttaaag  420

gata(--190nt-)tctagagtgttttccaccatctttcaaaggaaacatgtagtgtacc  680

ttcgaatgaaatggatttgtattaaacttttgccttagttattagggtctttctaattt  740

ttgattaacatatttttttaatttgtggtgtttatttctgtttttattaacaaacgaact  800

                                        GlyLysTyrIleThrPheIleProGlyPro
catatgctcctctctctttttttttttttctGGAAAGTACATAACATTTATACCTGGACCC  860

PheGlnTyrGlnLeuValIleLysAsnHisLysAsnLeuSerPheArgValSerSerAsn
TTCCAGTACCAGCTAGTGATAAAAAATCATAAACAATTAAGCTTTCGAGTATCTTCTAAT  920

LysLeuSerGluLysAlaPheSerIleSerGlnLeuGlnLysAsnHisAspMetAsnAsp
AAATTATCAGAGAAAGCATTTAGCATCAGTCAGTTGCAAAAAAATCATGACATGAATGAT  980

GluLeuHisLeuCysTyrGlnPheLeuGlnGLyLysSerLeuAspIleLeuLeuGluSer
GAGCTTCACCTCTGCTATCAATTTTTGCAAGGCAAATCTTTGGACATACTCTTGGAAAGC 1040
                                                    exon 3
MetMetGlnAsnLeuLysGlnValPheGluProGlnLeuLeuLysThrThrSerTrpAsp
ATGATGCAGAATCTAAAACAAGTTTTTGAACCCCAGCTGTTAAAAACCACAAGTTGGGAC 1100
```

SequID No: 3.(contd.2)

ThrAlaGluLeuTyrProPheCysSerSerIleIlePheGluIleThrPheThrThrIle
ACGGCAGAACTGTATCCATTCTGCAGCTCAATAATATTTGAGATCACATTTACAACTATA 1160

TyrGlyLysValIleValCysAspAsnAsnLysPheIleSerGluLeuArgAspAspPhe
TATGGAAAAGTTATTGTTTGTGACAACAACAAATTTATTAGTGAGCTAAGAGATGATTTT 1220

LeuLysPheAspAspLysPheAlaTyrLeuValSerAsnIleProIleGluLeuLeyGly
TTAAAATTTGATGACAAGTTTGCATATTTAGTATCCAACATACCCATTGAGCTTCTAGGA 1280

AsnValLysSerIleArgGluKysIleIleLysCysPheSerSerGluLysLeuAlaLys
AATGTCAAGTCTATTAGAGAGAAAATTATAAAATGCTTCTCATCAGAAAAGTTAGCCAAG 1340

MetGlnGlyTrpSerGluValPheGlnSerArgGlnAspAspLeuGluLysTyrTyrVal
ATGCAAGGATGGTCAGAAGTTTTTCAAAGCAGGCAAGATGACCTGGAGAAATATTATGTG 1400

HisGluAspLeuGluIleGlyA-
CACGAGGACCTTGAAATAGGAGgtaagaacttctgaatgagcacttgcctaaataaaaat 1460

catttacatagacctctgaaataaaaaaagacaaaatggcgaccttgaaaatttttttat 1520

gctctttctaattggctaatgataaatgtttactctgatataacctctataattgatatt 1580

ttttttttgctgaggtggtaaacagatacttaatggtgataatgagaaagcgtataact 1640
intron 3

aagctgcatttatccctcttatctcatccccgaccacaccgccccccccatacacattac 1700

attttaaactattctcattaagcagaaaattagacttcagaagcctattggttctcatta 1760

gcatgcagtgatccttggctggtctgtgtcctaacatcttttaattagcacactgcaaat 1820

SequID No: 3 (contd.3)

```
                                                        -laHisHis
ctaatcagtgtaataaacgctattaatcttcctttacacttattttctcccaCACATCAT 1880

PheGlyPheLeuTrpValSerValAlaSerThrIleProThrMetPheTrpAlaThrTyr
TTAGGCTTTCTCTGGGCCTCTGTGGCAAACACTATTCCAACTATGTTCTGGGCAACGTAT 1940
                                                        exon 4
TyrLeuLeuArgHisProGluAlaMetAlaAlaValArgAspGluIleAspArgLeuLeu
TATCTTCTGCGGCACCCAGAAGCTATGGCAGCAGTGCGTGACGAAATTGACCGTTTGCTG 2000

GlnSerThrGlyGlnLysGluGlySerGlyPheProIleHisLeuThrArgGluGlnLeu
CAGTCAACAGGTCAAAAGGAAGGGTCTGGATTTCCCATCCACCTCACCAGAGAACAATTG 2060

AspSerLeuIleCysLeu
GACAGCCTAATCTGCCTAGgtaattattttatctgttatgaagaaagaaggtacctctct 2120

gcaaactcggtttatcactcatagctgtttacaagaggtagaggacacagctgctaattg 2180

acataataactcccatttacatcaattataaattatgtagtttatagccgtagatcatct 2240
                                                        intron 4
cattgcatgtaaacataaggcctaxgtaattaactgtgxaaxgtatgxaaaaxxctaacc 2300

aaagctt(--550nt-)cctgactgaacttcttactgccaaagttaaattccataccaat 2960

gagttattctctattctctctgtattgacatttcatctgcggtatcctttagggtacaat 3020

attccaagtttctttagacaaacgcaggaacaaatgttcacatatttctgtttctttatt 3080

cctttgacaagtaggcgagcattttagcctatgttggtctcaaaaaaaatcttttaaata 3140

tgttccaggttctttaatgggacctttcaggagcaaaagtcctcccaggtttggtcaatg 3200

ttcaccctcxgtggccattgaggaaaatgcccxxxxxgttctagagattgttctcacttc 3260
```

SequID No: 3 (contd.4)

tcaggctaaggcccattgagcaatgccagaaagcatgccttatactagcagtcaatttgg 3320

aagtttgtagtttgtgtctttagcataggttatcaaataaattttatatttxcttttaaa 3380

aaaatctcaacattactaaaatacaaatatccttttatttttctttgcagaattatcggg 3440

gaacaaatccagaaaatttgtgtaaatttcgggtagttgctccacttgatacacagtatt 3500

tctgcatattgtaatttctatgaagatctaggttgcatttcccatacattcaagcagttt 3560

ccattgcatttttatgaataagatgacgcatactgggaagtaaggcaaatacactaaaag 3620

gaatatgtgtttgtattctgtatagttattactcttaaaaaaagtagttgtaattcatcc 3680

actcttttttactttcaacttttgctattaaaaaatcatttttaaatttcagtattaaag 3740

cagaaacatttaaatttattagaccagaaaaataacagattctagaactataatttgaat 3800

ccatttaagcccatagctagagctagagattttcactattggatcc 3846

CLAIMS

1. A DNA molecule selected from the following:

(a) DNA molecules containing the coding sequence set forth in SEQ Id No: 1 beginning at nucleotide 22 and ending at nucleotide 1541,

(b) DNA molecules containing the coding sequence set forth in SEQ Id No: 2 beginning at nucleotide 1 and ending at nucleotide 1242,

(c) DNA molecules capable of hybridizing with the DNA molecule defined in (a) or (b) under standard hybridization conditions defined as 2 x SSC at 65°C.

(d) cytochrome P450-encoding DNA molecules capable of hybridizing with the DNA molecule defined in (a), (b) or (c) under reduced stringency hybridization conditions defined as 6 x SSC at 55°C.

2. A DNA molecule according to Claim 1 (c) or (d) comprising an Hct-1 gene-associated sequence of another vertebrate species, especially a mammalian species and in particular a human Hct-1 gene-associated sequence,

3. A DNA molecule according to Claim 2 selected from the following:

(e) DNA molecules comprising one or more sequences selected from
(i) the sequence designated "intron 2" in SEQ Id No 3,
(ii) the sequence designated "exon 3" in SEQ Id No 3,
(iii) the sequence designated "intron 3" in SEQ Id No 3,
(iv) the sequence designated "exon 4" in SEQ Id No 3, and
(v) the sequence designated "intron 5" in SEQ Id No 3; and

(f) DNA molecules capable of hybridizing with the DNA molecules defined in (e) under standard hybridization conditions defined as 2 x SSC at 65°C.

(g) cytochrome P450-encoding DNA molecules capable of hybridizing with the DNA molecule defined in (e) or (f) under reduced stringency hybridization conditions defined as 6 x SSC at 55°C.

4. A DNA molecule comprising a human Hct-1 gene-associated sequence and selected from the following:

(e) DNA molecules comprising one or more sequences selected from

(i) the sequence designated "intron 2" in SEQ Id No 3,

(ii) the sequence designated "exon 3" in SEQ Id No 3,

(iii) the sequence designated "intron 3" in SEQ Id No 3,

(iv) the sequence designated "exon 4" in SEQ Id No 3, and

(v) the sequence designated "intron 5" in SEQ Id No 3; and

(f) DNA molecules capable of hybridizing with the DNA molecules defined in (e) under standard hybridization conditions defined as 2 x SSC at 65°C.

(g) cytochrome P450-encoding DNA molecules capable of hybridizing with the DNA molecule defined in (e) or (f) under reduced stringency hybridization conditions defined as 6 x SSC at 55°C.

5. A DNA molecule comprising a human Hct-1 gene-associated sequence and selected from the following:

(h) DNA molecules comprising contiguous pairs of sequences selected from

(i) the sequence designated "intron 2" in SEQ Id No 3,
(ii) the sequence designated "exon 3" in SEQ Id No 3,
(iii) the sequence designated "intron 3" in SEQ Id No 3,
(iv) the sequence designated "exon 4" in SEQ Id No 3, and
(v) the sequence designated "intron 5" in SEQ Id No 3; and

(i) DNA molecules capable of hybridizing with the DNA molecules defined in (h) under standard hybridization conditions defined as 2 x SSC at 65°C.

(j) cytochrome P450-encoding DNA molecules capable of hybridizing with the DNA molecule defined in (h) or (i) under reduced stringency hybridization conditions defined as 6 x SSC at 55°C.

6. A DNA molecule comprising a human Hct-1 gene-associated coding sequence selected from the following:

(k) DNA molecules comprising a contiguous coding sequence consisting of the sequences "exon 3" and "exon 4" in SEQ Id No 3, and

(l) DNA molecules capable of hybridizing with the DNA molecules defined in (k) under standard hybridization conditions defined as 2 x SSC at 65°C.

(m) cytochrome P450-encoding DNA molecules capable of hybridizing with the DNA molecule defined in (k) or (l) under reduced stringency hybridization conditions defined as 6 x SSC at 55°C.

7. A DNA molecule encoding an Hct-1 gene-associated coding sequence coded for by a DNA molecule as claimed in any of Claims 1 to 6, but which differs in sequence from the sequences of the DNA molecules claimed in Claims 1 to 6 by virtue of one or more amino acids of said Hct-1 gene-associated sequences being encoded by degenerate codons.

8. A DNA molecule consisting of a contiguous sequence of at least 18 nucleotides from the DNA sequence set forth in SEQ Id Nos: 1, 2 and 3.

9. A DNA sequence according to Claim 8 containing at least 24 and most preferably at least 30 nucleotide taken from said sequence.

10. The use of a DNA molecule according to Claim 8 or Claim 9 as a hybridization probe for isolating or detecting members of gene families and homologous DNA sequences related to the Hct-1 gene, especially a human gene sequence.

11 The use of a DNA molecule according to Claim 8 or Claim 9 in the diagnosis of neuropsychiatric disorders, endocrine disorders, immunological disorders, diseases of cognitive function or neurodegenerative diseases.

12. The use of a short (e.g. 10 to 25) oligonucleotide primer, capable of hybridising with a DNA molecule claimed in any of Claims 1 to 9 in the polymerase chain reaction (PCR) amplification of a genomic or cDNA from a biological sample for the purpose of diagnosis of neuropsychiatric disorders, endocrine disorders, immunological disorders, diseases of cognitive function or neurodegenerative diseases.

13. A cytochrome P450 protein, at least a portion of which is encoded by a DNA molecule as claimed in any of claims 1 to 7.

14. A protein selected from the following:

(i) the protein designated rat Hct-1 comprising the amino acid sequence set forth in SEQ Id No: 1 or a protein having substantial homology thereto,

(ii) the protein designated mouse Hct-1 comprising the amino acid sequence set forth in SEQ Id No: 2 or a protein having substantial homology thereto, or

(iii) the protein designated human Hct-1 comprising the amino acid sequence set forth in SEQ Id No: 3 or a protein having substantial homology thereto.

15. A protein according to Claim 14 having a degree of homology such that at least 50%, preferably at least 60% and most preferably at least 70% of the amino acids match said Seq.ID No: 1, 2 or 3).

16. A process for producing a Hct-1 polypeptide, which comprises culturing a transformed host and recovering the desired Hct-1 polypeptide, characterised in that the host is transformed with nucleic acid comprising a coding sequence as defined in any of Claims 1 to 7.

17. A process according to Claim 15 wherein the transformed host cell is a yeast, bacterial, insect or mammalian cell.

18. A process according to Claim 16 or Claim 17 wherein the nucleic acid comprises an expression construct or an expression vector.

19. A process according to Claim 18 wherein the vector is a vaccinia virus or baculovirus vector, a yeast plasmid or integration vector.

20. An antibody, especially a monoclonal antibody which binds to a Hct-1-pr tein.

21. The use of an antibody according to Claim 19 in the diagnosis of neuropsychiatric disorders, endocrine disorders, immunological disorders, diseases of cognitive function, neurodegenerative diseases or diseases of cognitive function.

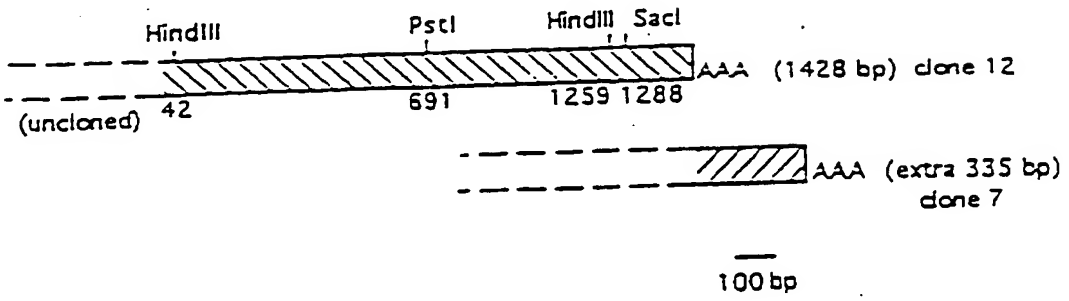
22. The use of a protein according to any of Claims 13 to 15, or an antibody according to Claim 18 in the design and/or manufacture of an antagonist to Hct-1 protein.

23. The use of a protein according to any of Claims 13 to 15, to effect a catalytic transformation of a substrate.

24. The use according to Claim 23 wherein the substrate is a steroid.

25. A transformed substrate when produced as a result of the use claimed in Claim 23 or Claim 24.

(a)



(b)

```
A  L  E  Y  Q  Y  V  M  K  N  P  K  Q  L  S  F  E  K  F  S
GCCTTGGAGTACCAGTATGTAATGAAAAACCCAAAACAATTAAGCTTTGAGAAGTTCAGC    60
R  R  L  S  A  K  A  F  S  V  K  K  L  L  T  N  D  D  L  S
CGAAGATTATCAGCGAAAGCCTTCTCTGTCAAGAAGCTGCTAACTAATGACGACCTTAGC   120
N  D  I  H  R  G  Y  L  L  Q  G  K  S  L  D  G  L  L  E
AATGACATTCACAGAGGCTATCTTCTTTTACAAGGCAAATCTCTTGGATGGTCTTCTGGAA  180
T  M  I  Q  E  V  K  E  I  F  E  S  R  L  L  K  L  T  D  W
ACCATGATCCAAGAAGTAAAAGAAATATTTGAGTCCAGACTGCTAAAACTCACAGATTGG   240
N  T  A  R  V  F  D  F  C  S  S  L  V  F  E  I  T  F  T  T
AATACAGCAAGAGTATTTGATTTCTGTAGTTCACTGGTATTTGAAATCACATTTACAACT   300
I  Y  G  K  I  L  A  A  N  K  K  Q  I  I  S  E  L  R  D  D
ATATATGGAAAAATTCTTGCTGCTAACAAAAAACAAATTATCAGTGAGCTGAGGGATGAT   360
F  L  K  F  D  D  H  F  P  Y  L  V  S  D  I  P  I  Q  L  L
TTTTTAAAATTTGATGACCATTTCCCATACTTAGTATCTGACATACCTATTCAGGTTCTA   420
R  N  A  E  F  M  Q  K  K  I  I  K  C  L  T  P  E  K  V  A
AGAAATGCAGAATTTATGCAGAAGAAAATTATAAAATGTCTCACACCAGAAAAAGTAGCT   480
Q  M  Q  R  R  S  E  I  V  Q  E  R  Q  E  M  L  K  K  Y  Y
CAGATGCAAAGACGGTCAGAAATTGTTCAGGAGAGGCAGGAGATGCTGAAAAAATACTAC   560
G  H  E  E  F  E  I  G  A  H  H  L  G  L  L  W  A  S  L  A
GGGCATGAAGAGTTTGAAATAGGAGCACATCATCTTGGCTTGCTCTGGGCCTCTCTAGCA   600
N  T  I  P  A  M  F  W  A  M  Y  Y  L  L  Q  H  P  E  A  M
AACACCATTCCAGCTATGTTCTGGGCAATGTATTATCTTCTTCAGCATCCAGAAGCTATG   660
E  V  L  R  D  E  I  D  S  F  L  Q  S  T  G  Q  K  K  G  P
GAAGTCCTGCGTGACGAAATTGACAGCTTCCTGCAGTCAACAGGTCAAAAGAAAGGACCT   720
G  I  S  V  H  F  T  R  E  Q  L  D  S  L  V  C  L  E  S  A
GGAATTTCTGTCCACTTCACCAGAGAACAATTGGACAGCTTGGTCTGCCTGGAAAGCGCT   780
I  L  E  V  L  R  L  C  S  Y  S  S  I  I  R  E  V  Q  E  D
ATTCTTGAGGTTCTGAGGTTGTGCTCCTACTCCAGCATCATCCGTGAAGTGCAAGAGGAT   840
M  D  F  S  S  E  S  R  S  Y  R  L  R  K  G  D  F  V  A  V
ATGGATTTCAGCTCAGAGAGTAGGAGCTACCGTCTGCGGAAAGGAGACTTTGTAGCTGTC   900
F  P  P  M  I  H  N  D  P  E  V  F  D  A  P  K  D  F  R  F
TTTCCTCCAATGATACACAATGACCCAGAAGTCTTCGATGCTCCAAAGGACTTTAGGTTT   960
D  R  F  V  E  D  G  K  K  K  T  T  F  F  K  G  G  K  K  L
GATCGCTTCGTAGAAGATGGTAAGAAGAAAACAACGTTTTTCAAAGGAGGAAAAAAGCTG  1020
K  S  Y  I  I  P  F  G  L  G  T  S  K  C  P  G  R  Y  F  A
AAGAGTTACATTATACCATTTGGACTTGGAACAAGCAAATGTCCAGGCAGATACTTTGCA  1080
I  N  E  M  K  L  L  V  I  I  L  L  T  Y  F  D  L  E  V  I
ATTAATGAAATGAAGCTACTAGTGATTATACTTTTAACTTATTTTGATTTAGAAGTCATT  1140
D  T  K  P  I  G  L  N  H  S  R  M  F  L  G  I  Q  H  P  D
GACACTAAGCCTATAGGACTAAACCACAGTCGCATGTTTCTGGGCATTCAGCATCCAGAC  1200
S  D  I  S  F  R  Y  K  A  K  S  W  R  S  ***
TCTGACATCTCATTTAGGTACAAGGCAAAATCTTGGAGATCCTGAAAGGGTGGCAGAGAA  1260

GCTTAGCGGAATAAGGCTGCACATGCTGAGCTCTGTGATTTGCTGTACTCCCCAAATGCA  1320

GCCACTATTCTTGTTTGTTAGAAAATGGCAAATTTTTATTTGATTGCGATCCATCCAGTT  1380

TGTTTTGGGTCACAAAACCTGTCATAAAATTTGGCGCTGTCATGGTGTaaaaaaatgtca  1440

tggcaatcatttcaggataaggtaaaataacgttttcaagtttgtacttactatgatttt  1500

tatcacttgtagtgaatgtgcttttccagtaataaatttgcgccagggtgatttttttta  1560

attactgaaatcctctaatatcggttttatgtgctgccagaaaagtgtgccatcaatgga  1620

cagtataacaatttccagttttccagagaagggagaaacttaagccccatgagttacgctg  1680

tataaaattgttctcttcaacttataatatcaataatgtcctatatcaccaggttacctttg  1740

cattaaatcgagttttgcaaaag  1763
```

FIG 1

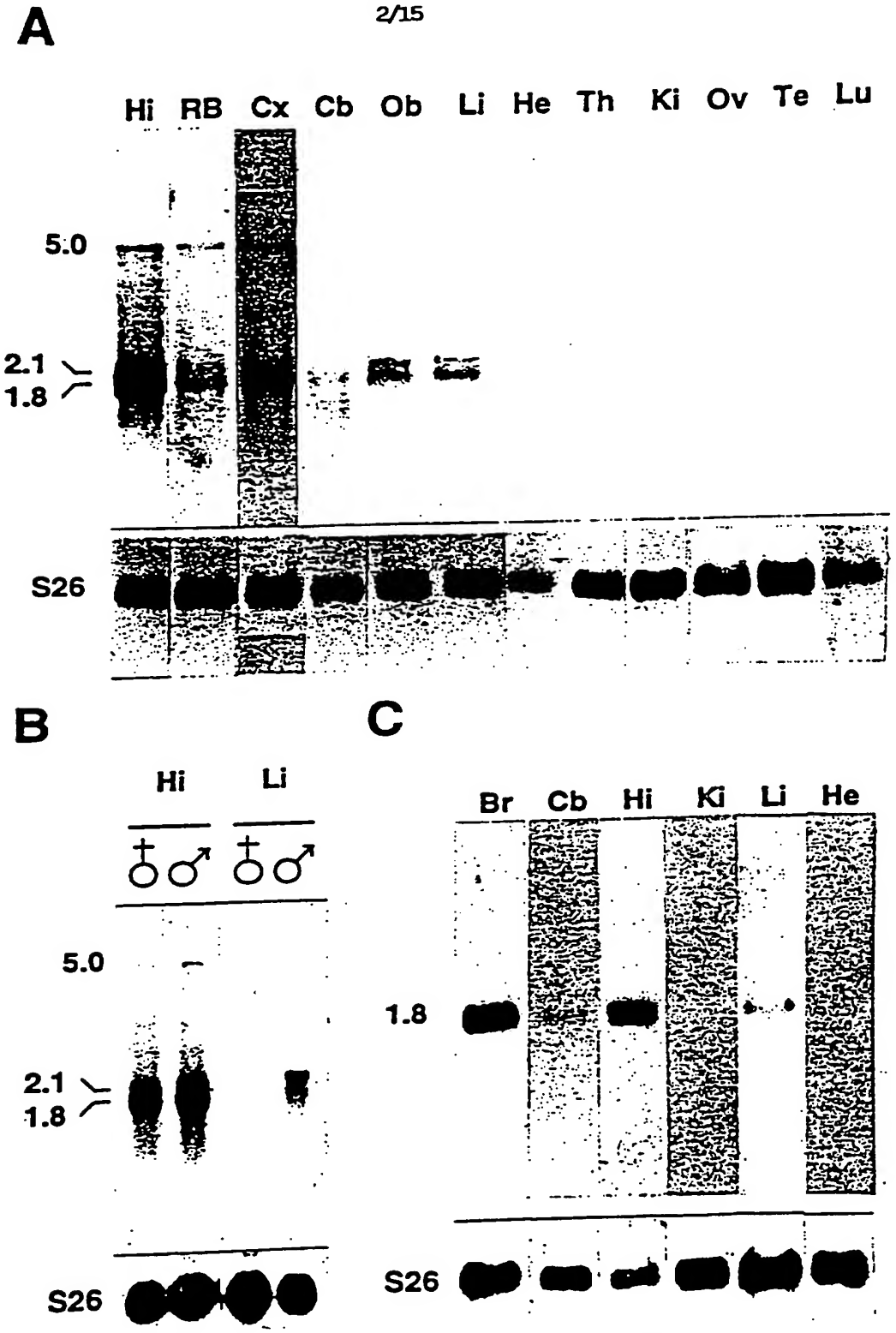
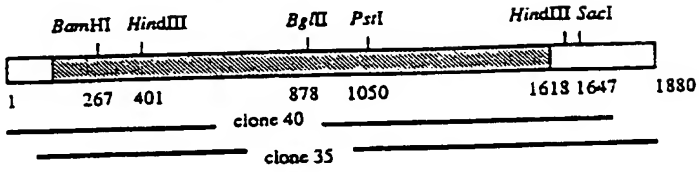


FIG 2




```
ggcaggcacagccctctggtctaagaagagagggcactgtgcagaagccatcgctcccctaC   60
                    M  Q  G  A  T  T  L  D  A  A  S  P  G  P     14
AGAGCCGCCAGCTCGTCGGGATGCAGGGAGCCACGACCCTAGATGCCGCCTCGCCAGGGC   120
  L  A  L  L  G  L  L  F  A  A  T  L  L  L  S  A  L  F  L  L    34
CTCTCGCCCTCCTAGGCCTTCTCTTTGCCGCCACCTTACTGCTCTCGGCCCTGTTCCTCC   180
  T  R  R  T  R  R  P  R  E  P  P  L  I  K  G  W  L  P  Y  L    54
TCACCCGGGCGCACCAGGCGCCCTCGTGAACCACCCTTGATAAAAGGTTGGCTTCCTTATC   240
  G  M  A  L  K  F  F  K  D  P  L  T  F  L  K  T  L  Q  R  Q    74
TTGGCATGGCCCCTGAAATTCTTTAAGGATCCGTTAACTTTTTTGAAAACTCTTCAAGGC   300
  H  G  D  T  F  T  V  F  L  V  G  K  Y  I  T  F  V  L  N  P    94
AACATGGTGACACTTTCACTGTCTTCCTTGTGGGGAAGTATATAACATTTGTTCTGAACC   360
  F  Q  Y  Q  Y  V  T  K  N  P  K  Q  L  S  F  Q  K  F  S  E   114
CTTTCCAGTACCAGTATGTAACGAAAAACCCAAAACAATTAAGCTTTCAGAAGTTCAGCA   420
  R  L  S  A  K  A  F  S  V  K  K  L  L  T  D  D  D  L  N  E   134
GCCGATTATCAGCGAAAGCCTTCTCTGTAAAGAAGCTGCTACTGATGACGACCTTAATG   480
  D  V  H  R  A  Y  L  L  L  Q  G  K  P  L  D  A  L  L  E  T   154
AAGACGTTCACAGAGCCTATCTACTTCTACAAGGCAAACCTTTGGATGCTCTTCTGGAAA   540
  M  I  Q  E  V  K  E  L  F  E  S  Q  L  L  K  I  T  D  W  N   174
CTATGATCCAAGAAGTAAAAGAATTATTTGAGTCCCAACTGCTAAAAATCACAGATTGGA   600
  T  E  R  I  F  A  F  C  G  S  L  V  F  E  I  T  F  A  T  L   194
ACACAGAAAGAATATTTGCATTCTGTGGCTCACTGGTATTTGAGATCACATTTGCGACTC   660
  Y  G  K  I  L  A  G  N  K  K  Q  I  I  S  E  L  R  D  D  F   214
TATATGGAAAAATTCTTGCTGGTAACAAGAAACAAATTATCAGTGAGCTAAGGGATGATT   720
  F  K  F  D  D  M  F  P  Y  L  V  S  D  I  P  I  Q  L  L  R   234
TTTTTAAATTTGATGACATGTTCCCATACTTAGTATCTGACATACCTATTCAGCTTCTAA   780
  N  E  E  S  M  Q  K  K  I  I  K  C  L  T  S  E  K  V  A  Q   254
GAAATGAAGAATCTATGCAGAAGAAAATTATAAAATGCCTCACATCAGAAAAAGTAGCTC   840
  M  Q  G  Q  S  K  I  V  Q  E  S  Q  D  L  L  K  R  Y  Y  R   274
AGATGCAAGGACAGTCAAAAATTGTTCAGGAAAGCCAAGATCTGCTGAAAAGATACTATA   900
  H  D  D  P  E  I  G  A  H  H  L  G  F  L  W  A  S  L  A  N   294
GGCATGACGATTCTGAAATAGGAGCACATCATCTTGGCTTTCTCTGGGCCTCTCTAGCAA   960
  T  I  P  A  M  F  W  A  M  Y  Y  I  L  R  H  P  E  A  M  E   314
ACACCATTCCAGCTATGTTCTGGGCAATGTATTATATTCTTCGGCATCCTGAAGCTATGG  1020
  A  L  R  D  E  I  D  S  F  L  Q  S  T  G  Q  K  K  G  P  G   334
AAGCCCTGCGTGACGAAATTGACAGTTTCCTGCAGTCAACAGGTCAAAAGAAAGGGCCTG  1080
  I  S  Y  H  F  T  R  E  Q  L  D  S  L  V  C  L  E  S  T  I   354
GAATTTCAGTCCACTTCACCAGAGAACAATTGGACAGCTTGGTCTGCCTGGAAAGCACTA  1140
  L  E  V  L  R  L  C  S  Y  S  S  I  I  R  E  V  Q  E  D  M   374
TTCTTGAGGTTCTGAGGCTGTGCTCATACTCCAGCATCATCCGAGAAGTGCAGGAGGATA  1200
  N  L  S  L  E  S  K  S  F  S  L  R  K  G  D  F  V  A  L  F   394
TGAATCTCAGCTTAGAGAGTAAGAGTTTCTCTCTGCGGAAAGGAGATTTTGTAGCCTCT  1260
  P  P  L  I  H  N  D  P  E  I  F  D  A  P  K  E  F  R  F  D   414
TTCCTCCACTCATACACAATGACCCGGAAATCTTCGATGCTCCAAAGGAATTTAGGTTCG  1320
  R  F  I  E  D  G  K  K  K  S  T  F  F  K  G  G  K  R  L  K   434
ATCGGTTCATAGAAGATGGTAAGAAGAAAAGCACGTTTTTCAAAGGAGGGAAGAGGCTGA  1380
  T  Y  V  M  P  F  G  L  G  T  S  K  C  P  G  R  Y  F  A  V   454
AGACTTACGTTATGCCTTTTGGACTCGGAACAAGCAAATGTCCAGGGAGATATTTTGCAG  1440
  N  E  M  K  L  L  L  I  E  L  L  T  Y  F  D  L  E  I  I  D   474
TGAACGAAATGAAGCTACTGCTGATTGAGCTTTTAACTTATTTTGATTTAGAAATTATCG  1500
  R  K  P  I  G  L  N  H  S  R  M  F  L  G  I  Q  H  P  D  S   494
ACAGGAAGCCTATAGGGCTAAATCACAGTCGGATGTTTTTAGGTATTCAGCACCCCGATT  1560
                                                               507
  A  V  S  F  R  Y  K  A  K  S  W  R  S  ***
CTGCCGTCTCCTTTAGGTACAAAGCAAAATCTTTGGAGAAGCTGAAAGTGTGGCAGAGAAG  1620

CTTTGCAGAGTAAGGCTGCATGTGCTGAGCTCCGTGATTTGGTGCACTCCCCCAAATGCA  1680

ACCGCTACTCTTGTTTGAAAATGGCAAATTTATATTTGGTTGAGATCAATCCAGTTGGTT  1740

TTGGGTCACAAAACCTGTCATAAAATAAAGCAGTGTGATGGcttaaaaaatgtcatggca  1800

atcatttcaggataaggtaaaatcaacatttcaagtttgtacttactatgatttttatca  1860

ttcgtagtgaatgtgctttt  1880
```

FIG 3

A

```
MoHct-1     1 MQGATTLDAASPGPLALLGLLFAATLLLSALFLLTRRTRRPREPPLIKGW  50
HuCYP7      1 M--MTTSLIWGIAIAACCCL----------WLILGIRRRQTG-EPPLENGL  38

 51 GWLPYLGMALKFFKDPLTFLKTLQRQHGDTFTVFLVGKYITFVLNPFQYQYVTKNPKQLSFQKF 112
 39 GLIPYLGCALQFGANPLEFLRANQRKHGHVFTCKLMGKYVHFITNPLSYHKVLCHGKYFDWKKF 100

113 SSRLSAKAFSVKKLLT-DDDLNEDVHRAYL-LLQGKPLDALLETMI---QEVKELFESQLLKIT 171
101 HFATSAKAFGHRSIDPMDGNTTENINDTFIKTLQGHALNSLTESMMENLQRIMRPPVSSNSKTA 164

172 DWNTERIFAFCGSLVFEITFATLYGKILA---GNKKQIISELRDDFFKFDDM-FPYLVSDIPIQ 231
165 AWVTEGMYSFCYRVMFEAGYLTIFGRDLTRRDTQKAHILNNL--DNFKQFDKVFPALVAGLPIH 226

240 LLRNEESMQKKIIKCLTSEKVAQMQGQSKIVQESQDLLKRYYRHDDPEIGAHHLGFLWASLANT 295
227 MFRTAHNAREKLAESLRHENLQKRESISELISLRMFLNDTLSTFDDLEKAKTHLVVLWASQANT 290

296 IPAMFWAMYYILRHPEAMEALRDEIDSFLQSTGQKKG-PGISVHFTREQLDSLVCLESTILEVL 358
291 IPATFWSLFQMIRNPEAMKAATEEVKRTLENAGQKVSLEGNPICLSQAELNDLPVLNSIIKESL 354

359 RLCSYSSIIREVQEDMNLSLESKSFSLRKGDFVALFPPLIHNDPEIFDAPKEFRFDRF-IEDGK 421
355 RLSSASLNIRTAKEDFTLHLEDGSYNIRKDSIIALYPQLMHLDPEIYPDPLTFKYDRYLDENGK 418

422 KKSTFFKGGKRLKTYVMPFGLGTSKCPGRYFAVNEMKLLLIELLTYFDLEIID--RKPIGLNHS 483
419 TKTTFYCNGLKLKYYYMPFGSGATICPGRLFAIHEIKQFLILMLSYFELELIEGQAKCPPLDQS 482

484 RMFLGIQHPDSAVSFRYKAKSWRS* 507
483 RAGLGILPPLNDIEFKYKFKHL* 504
```

B

```
MoHct-1   F G L G T S K C P G R Y F A
HuCYP7    F G S G A T I C P G R L F A
CYP17     F G A G P R S C V G E M L A
CYP11B    F G F G M R Q C L G R R L A
CYP21B    F G C G A R V C L G E P V A
CYP11A1   F G W G V R Q C L G R R I A
CYP27     F G Y G V R A C L G R R I A
```

C

```
V C L E S T I L E V L R L C S
P V L N S I I K E S L R L S S
V L L E H T I R E V L R I R P
P L L R A A L K E T L R L Y P
P L L N A T I A E V L R L P V
P L L K A S I K E T L R L H P
P L L K A V L K E T L R L Y P
```

FIG 4

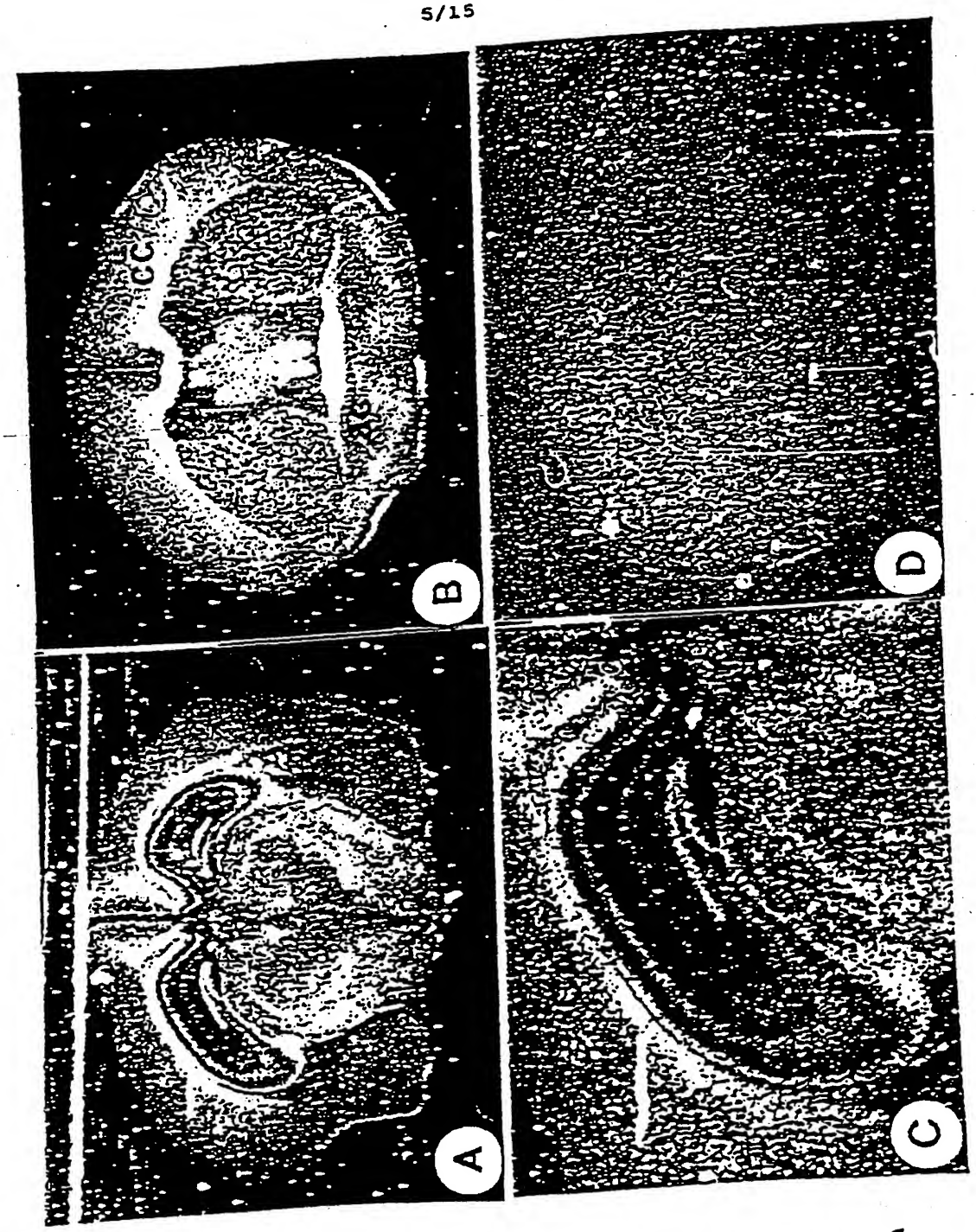


FIG 5

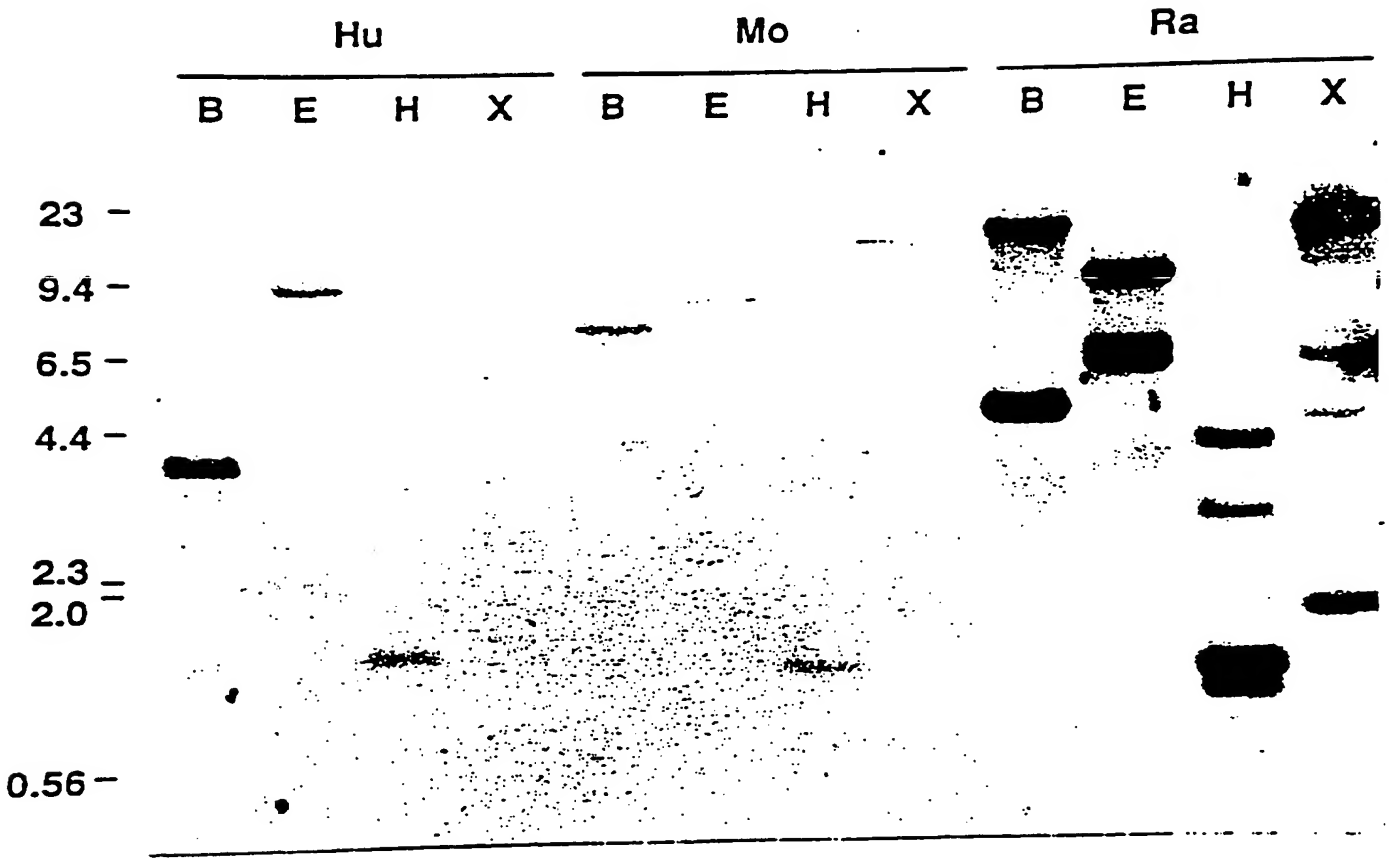


FIG 6

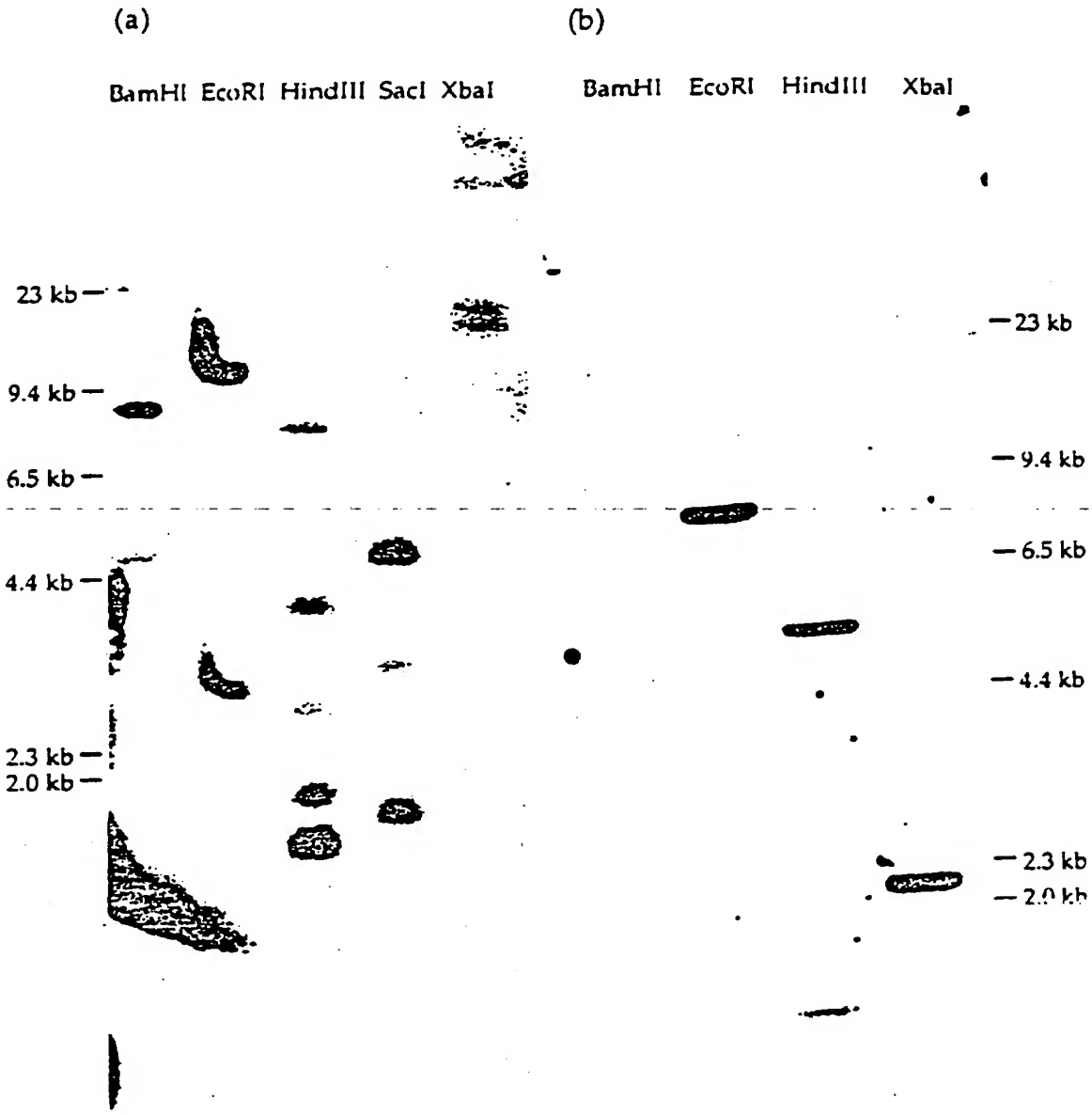


FIG 7

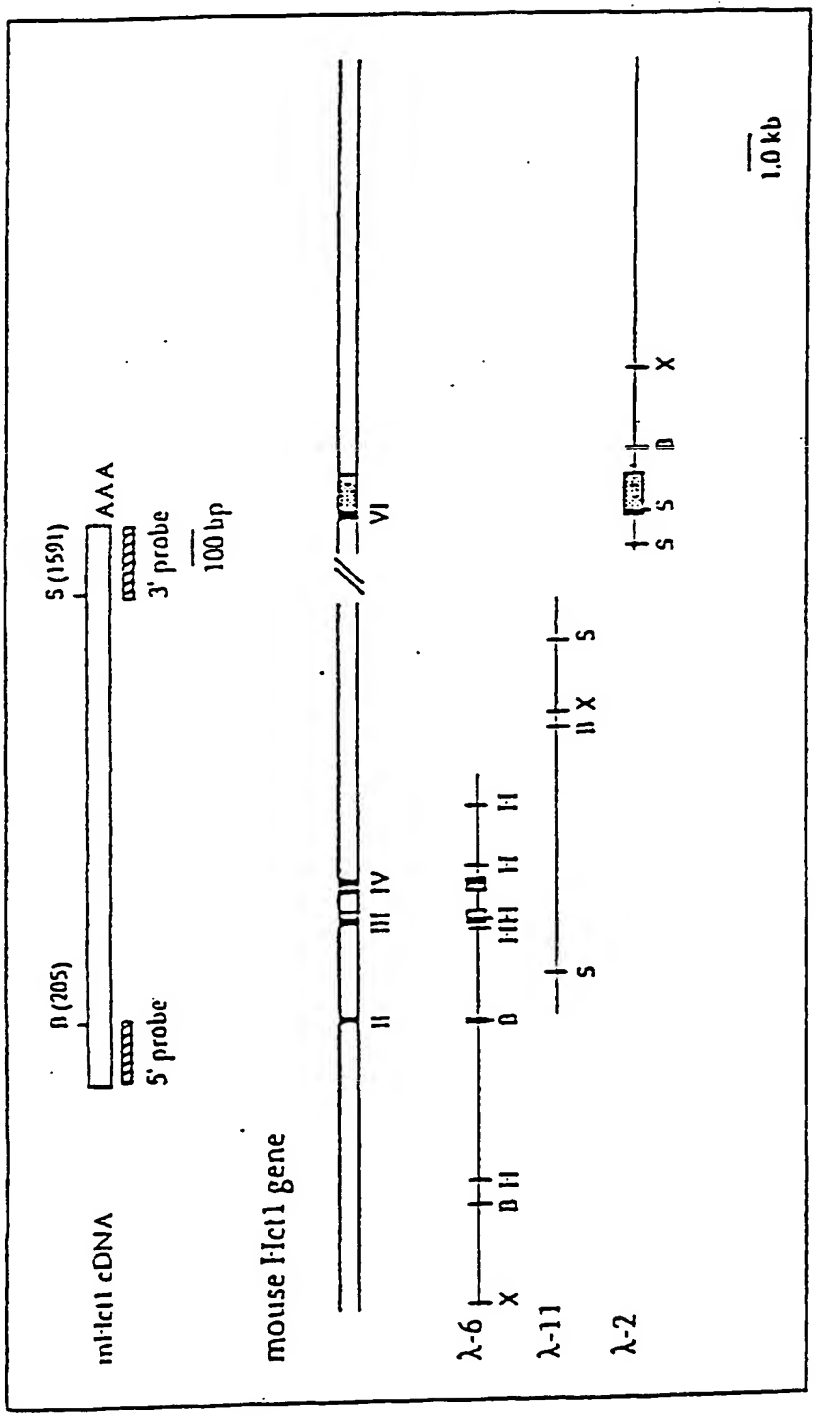


FIG 8

FIG 9

Partial nucleotide sequence of human genomic Hct-1 (CYP7B1) and the encoded polypeptide

```
ggatccaaccaagtttccagatcttataaatgtggtgaatggtgaatgacttcctgaaga      60
atggatgaatggatgtgttctagtttggaatcctgtgtcagtcacaagtcaatatgtgac     120
cttgaacatgttattaaatctcccacatccataaaagtgaaaatgctggcattagtggat     180
ttttgccagtgttgaattagacatttatttgtgagtaccttgctccatacagtatggtcat     240
ttatttgagttaaaacttgttgtatttgaacaaaacttcagatgacacctaagcatgaaaaa     300
                    intron 2
gctcttttatgaagtataaatacttcagaaatggaatggcatgttgccaatttgttttctgc     360
tttattgagggaaatatatgagaagtatttaagtcaggggattatgaggaatatttaaag      420
gata(--190nt-)tctagagtgtttttccaccatcttttcaaaggaaacatgtagtgtacc     680
ttcgaatgaaatggatttgtattaaacttttgccttagttattagggtcttttctaatttt     740
ttgattaacatatttttttaatttgtggtgtttatttctgtttttattaacaaacgaact      800
                                    GlyLysTyrIleThrPheIleProGlyPro
catatgctccttcttcttttttttttttttctGGAAAGTACATAACATTTATACCTGGACCC     860
PheGlnTyrGlnLeuValIleLysAsnHisLysAsnLeuSerPheArgValSerSerAsn
TTCCAGTACCAGCTAGTGATAAAAAATCATAAACAATTAAGCTTTCGAGTATCTTCTAAT      920
LysLeuSerGluLysAlaPheSerIleSerGlnLeuGlnLysAsnHisAspMetAsnAsp
AAATTATCAGAGAAAGCATTTAGCATCAGTCAGTTGCAAAAAAATCATGACATGAATGAT      980
GluLeuHisLeuCysTyrGlnPheLeuGlnGLyLysSerLeuAspIleLeuLeuGluSer
GAGCTTCACCTCTGCTATCAATTTTTGCAAGGCAAATCTTTGGACATACTCTTGGAAAGC     1040
                    exon 3
MetMetGlnAsnLeuLysGlnValPheGluProGlnLeuLeuLysThrThrSerTrpAsp
ATGATGCAGAATCTAAAACAAGTTTTTGAACCCCAGCTGTTAAAAACCACAAGTTGGGAC     1100
ThrAlaGluLeuTyrProPheCysSerSerIleIlePheGluIleThrPheThrThrIle
ACGGCAGAACTGTATCCATTCTGCAGCTCAATAATATTTGAGATCACATTTACAACTATA     1160
TyrGlyLysValIleValCysAspAsnAsnLysPheIleSerGluLeuArgAspAspPhe
TATGGAAAAGTTATTGTTTGTGACAACAACAAATTTATTAGTGAGCTAAGAGATGATTTT     1220
LeuLysPheAspAspLysPheAlaTyrLeuValSerAsnIleProIleGluLeuLeyGly
TTAAAATTTGATGACAAGTTTGCATATTTAGTATCCAACATACCCATTGAGCTTCTAGGA     1280
AsnValLysSerIleArgGluKysIleIleLysCysPheSerSerGluLysLeuAlaLys
AATGTCAAGTCTATTAGAGAGAAAATTATAAAATGCTTCTCATCAGAAAAGTTAGCCAAG     1340
MetGlnGlyTrpSerGluValPheGlnSerArgGlnAspAspLeuGluLysTyrTyrVal
ATGCAAGGATGGTCAGAAGTTTTTCAAAGCAGGCAAGATGACCTGGAGAAATATTATGTG     1400
```

FIG 9 (Contd page 2)

```
HisGluAspLeuGluIleGlyA-
CACGAGGACCTTGAAATAGGAGgtaagaacttctgaatgagcacttgcctaaataaaaat 1460

catttacatagaccttctgaaataaaaaaagacaaaatggcgaccttgaaaatttttttat 1520

gctcttttctaattggctaatgataaatgtttacttctgatataaccttctataattgatatt 1580

ttttttttgctgaggtggtaaacagatacttaatggtgataatgagaaagcgtataact 1640
                    intron 3
aagctgcatttatccctcttatctcatcccgaccacaccgcccccccccatacacattac 1700

attttaaactattctcattaagcagaaaattagacttcagaagcctattggttctcatta 1760

gcatgcagtgatccttggctggtctgtgtcctaacatcttttaattagcacactgcaaat 1820

                                                    -laHisHis
ctaatcagtgtaataaacgctattaatcttccttcacacttatttttctcccaCACATCAT 1880

PheGlyPheLeuTrpValSerValAlaSerThrIleProThrMetPheTrpAlaThrTyr
TTAGGCTTTCTCTGGGCCTCTGTGGCAAACACTATTCCAACTATGTTCTGGGCAACGTAT 1940
                    exon 4
TyrLeuLeuArgHisProGluAlaMetAlaAlaValArgAspGluIleAspArgLeuLeu
TATCTTCTGCGGCACCCAGAAGCTATGGCAGCAGTGCGTGACGAAATTGACCGTTTGCTG 2000

GlnSerThrGlyGlnLysGluGlySerGlyPheProIleHisLeuThrArgGluGlnLeu
CAGTCAACAGGTCAAAAGGAAGGGTCTGGATTTCCCATCCACCTCACCAGAGAACAATTG 2060

AspSerLeuIleCysLeu
GACAGCCTAATCTGCCTAGgtaattattttatctgttatgaagaaagaaggtacctcttct 2120

gcaaactcggtttatcactcatagctgtttacaagaggtagaggacacagcctgctaattg 2180

acataataactcccatttacatcaattataaattatgtagtttatagccgtagatcatct 2240
                    intron 4
cattgcatgtaaacataaggcctaxgtaattaactgtgxaaxgtatgxaaaaxxctaacc 2300

aaagctt(--550nt-)cctgactgaacttcttactgccaaagttaaattccataccaat 2960

gagttattctctattcctcctgtattgacatttcatctgcggtatcccttagggtacaat 3020

attccaagtttctttagacaaacgcaggaacaaatgttcacatatttctgtttctttatt 3080

cctttgacaagtaggcgagcattttagcctatgttggtctcaaaaaaaatcctttaaata 3140

tgttccaggttctttaatgggacctttcaggagcaaaagtcctcccaggtttggtcaatg 3200

ttcaccctcxgtggccattgaggaaaatgcccxxxxxgttctagagattgttctcacttc 3260

tcaggctaaggcccattgagcaatgccagaaagcatgccttatactagcagtcaatttgg 3320

aagtttgtagtttgtgtctttagcataggttatcaaataaatttttatatttxcttttaaa 3380

aaaatctcaacattactaaaatacaaatatccttttatttttcttgcagaattatcggg 3440

gaacaaatccagaaaatttgtgtaaatttcgggtagttgctccacttgatacacagtatt 3500
```

FIG 9 (Contd page 3)

tctgcatattgtaatttctatgaagatctaggttgcatttcccatacattcaagcagttt 3560

ccattgcatttttatgaataagatgacgcatactgggaagtaaggcaaatacactaaaag 3620

gaatatgtgtttgtattctgtatagttattacttcttaaaaaaagtagttgtaattcatcc 3680

actctttttactttcaacttttgctattaaaaaatcatttttaaatttcagtattaaag 3740

cagaaacatttaaatttattagaccagaaaaataacagattctagaactataatttgaat 3800

ccatttaagcccatagctagagctagagattttcactattggatcc 3846

Comparison of human and mouse Hct-1 (Cyp7b1) sequences (exons III and IV)

```
human
G K Y I T F I P G P F Q Y Q L V I K N H K Q L S F R V S S N
! ! ! ! ! ! .     ! ! ! ! !   ! ! !   ! ! ! ! !       ! 
G K Y I T F V L N P F Q Y Q Y V T K N P K Q L S F Q K F S S
mouse

K L S E K A F S I S Q L Q K N H D M N D E L H L C Y Q F L Q
. ! ! ! ! ! ! .     !       ! . ! . . !     ! !   ! !
R L S A K A F S V K K L L T D D D L N E D V H R A Y L L L Q

G K S L D I L L E S M M Q N L K Q V F E P Q L L K T T S W D
! !   ! !   ! ! ! . ! . !   ! .   ! !   ! ! ! !   ! !
G K P L D A L L E T M I Q E V K E L F E S Q L L K I T D W N

T A E L Y P F C S S I I F E I T F T T I Y G K V I V C D N N
!     . . ! !   ! . . ! ! ! ! !   ! . ! ! ! . .
T E R I F A F C G S L V F E I T F A T L Y G K I L A G N K K

K F I S E L R D D F L K F D D K F A Y L V S N I P I E L L G
  . ! ! ! ! ! ! !   ! ! ! !   ! . ! ! ! ! ! ! !   ! !
Q I I S E L R D D F F K F D D M F P Y L V S D I P I Q L L R

N V K S I R E K I I K C F S S E K L A K M Q G W S E V F Q S
!   ! .   ! ! ! ! !   . ! ! ! . ! ! ! ! ! . ! !
N E E S M Q K K I I K C L T S E K V A Q M Q G Q S K I V Q E

                    <--exon III - exon IV-->
R Q D D L E K Y Y V H E D L E I G - A H H F G F L W V S V A
  ! !   !   . ! !   ! . !   ! ! !   ! ! !   ! ! ! ! ! . !
S Q D L L K R Y Y R H D D S E I G - A H H L G F L W A S L A

S T I P T M F W A T Y Y L L R H P E A M A A V R D E I D R L
  ! ! !   ! ! ! !   ! ! . ! ! ! ! ! ! ! ! . ! ! ! ! ! !
N T I P A M F W A M Y Y I L R H P E A M E A L R D E I D S F

L Q S T G Q K E G S G F P I H L T R E Q L D S L I C L
! ! ! ! ! ! !   ! !   . !   ! ! ! ! ! ! ! ! . ! !
L Q S T G Q K K G P G I S V H F T R E Q L D S L V C L
```

Shared identity = 163/266 residues; 61% identity (74% over exon IV)

FIG 10

A. Kozak sequences in mRNAs for steroidogenic P450's

Nucleotide sequences conforming are in bold; sequences diverging from the consensus are underlined

consensus
yyRyy ATG R

(a)

BovCYP21	- CTCCAGCC	ATG GTCCTCG
HumCYP21	- GTCTCGCC	ATG CTGCTCC
HumCYP17	- CAGCCACC	ATG TGGGAGC
MusCYP7B	- TCGTCGGG	ATG CAGGGAG
HumCYP7	- TTTGCAAA	ATG ATGACCA
RatCYP7	- TTTGCAAA	ATG ATGACTA
MusCYP7	- TTTGCAAA	ATG ATGAGCA
RabCYP27	- TCGGATCC	ATG GCTGCGC
RatCYP27	- CACGATCT	ATG GCTGTGT

Sequence selected for the mouse Hct-1 coding sequence in vaccinia virus

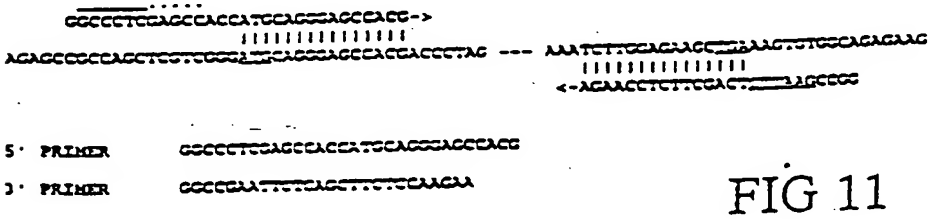
*MusCYP7B** - TCGCCACC ATG CAGGGAG

B. Mutagenesis of the 5' end of the mouse Hct-1 cDNA to create a near-consensus translation initiation region surrounding the ATG (AUG)

(b)

(i). Sequence surrounding the initiating ATG and the translation termination site

(ii). PCR primers for modification of the translation initiation site and 3' truncation of the cDNA clone



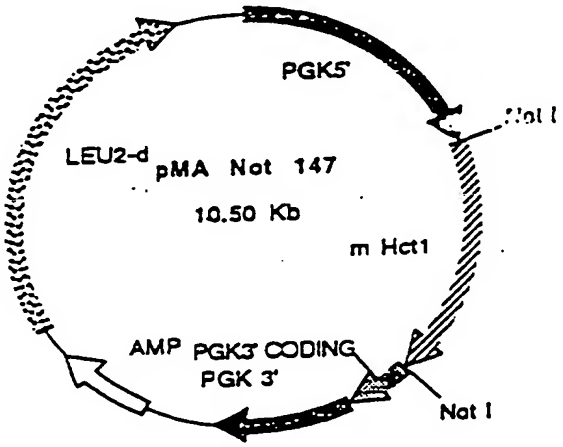
5' PRIMER

3' PRIMER

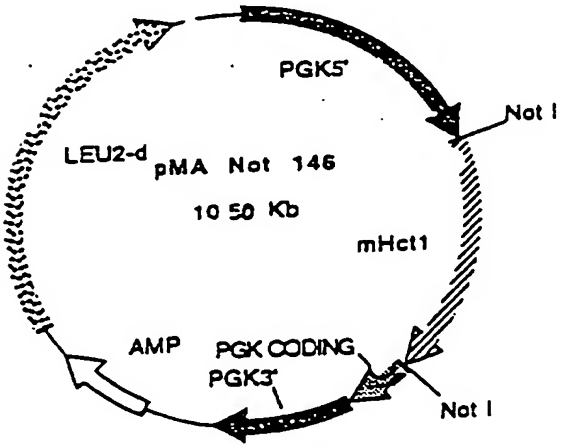
FIG 11

Yeast expression vectors containing the mouse Hct-1 coding sequence

A



B



Clone pMA Not 146 (A) contains the mouse Hct-1 cDNA clone 35; pMA Not 147 (B) contains cDNA clone 40.

FIG 12

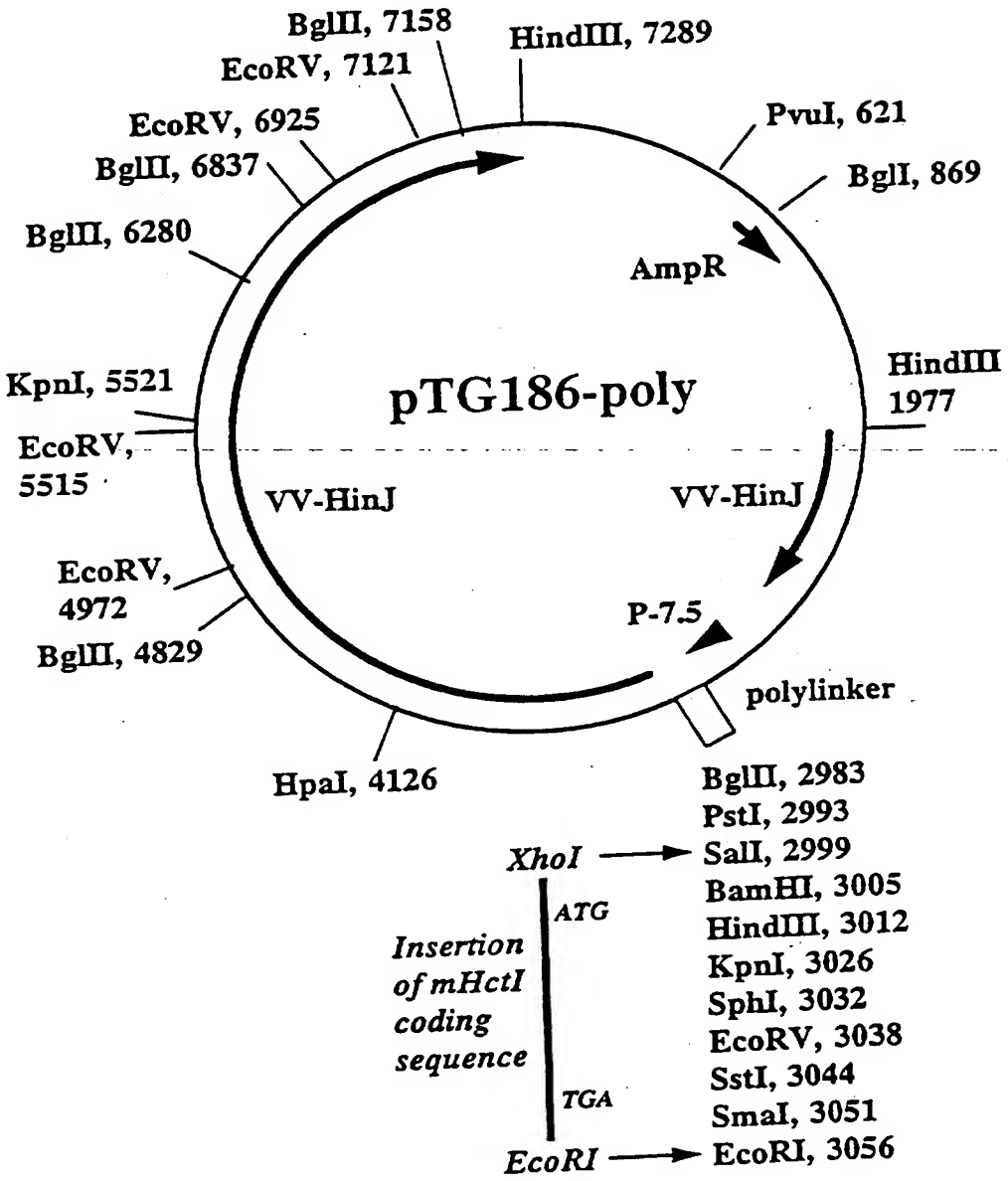


FIG 13

INTERNATIONAL SEARCH REPORT

Inte· onal Application No
PC[redacted] 95/02465

A. CLASSIFICATION OF SUBJECT MATTE[redacted]
IPC 6 C12N15/53 C12N9/02 A61K39/395 C12Q1/68 G01N33/577
//C12P21/02

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N A61K C12Q G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	FEBS LETT., vol. 268, no. 1, July 1990 pages 137-140, NISHORO ET AL. 'Molecular cloning and sequence analysis of cDNA encoding human cholesterol 7-alpha-hydroxylase' see the whole document ---	13-15, 17,18, 23-25
P,X	EP,A,0 648 840 (NORTHEASTERN OHIO UNIVERSITIES) 19 April 1995 see the whole document --- -/--	13-15, 17,18, 22-25

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
26 February 1996	15. 03. 96

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl, Fax (+ 31-70) 340-3016	Gac, G

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS SIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	FEBS LETTERS, vol. 257, no. 1, October 1989 pages 97-100, NOSHIRO ET AL. 'Molecular cloning for cholesterol 7-alpha-hydroxylase from rat liver microsomes' see the whole document	13,14
A		15,17, 18,20

X	BIOCHEM. BIOPHYS. RES. COMMUN., vol. 185, no. 2, 1992 pages 588-595, KARAM ET AL. 'Polymorphisms of human vholesterol 7-alpha-hydroxylase' see the whole document	13-15,23
A		1-12

X	PROC. NATL ACAD. SCI., vol. 83, June 1986 pages 4243-4247, CHUNG ET AL. 'Structure of a bovine gene for P450c21 (steroid 21-hydroxylase) defines a novel cytochrome P-450 gene family' see the whole document	13

A	BIOCHEM. PHARMACOLOGY, vol. 44, no. 6, 25 September 1992 pages 1223-1225, BAHMRE ET AL. 'Microsomal cytochrome P450 in human brain regions' see the whole document	11,14, 21,23

X	J. STEROID BIOCHEM., vol. 23, no. 6a, November 1985 pages 919-927, THERON ET AL. 'Evidence that estradiol-2/4-hydroxylase activities in rat hypothalamus and hippocampus differ qualitatively and involve multiple forms of P-450 : ontogenic and inhibition studies' see the whole document	25
A		14,22-24

A	NEUROSCIENCE, vol. 42, no. 1, 1991 pages 215-235, VOLK 'Mapping of phenytoin-inducible cytochrome p450 immunoreactivity in the mouse central nervous system' see pages 215, 221, 226 and 229	14,20, 23-25

-/--

3

INTERNATIONAL SEARCH REPORT

Intr ional Application No
PC 95/02465

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	BRAIN RES., vol. 536, no. 1-2, 17 December 1990 pages 339-343, ANANDATHEERTHAVARADA ET AL. 'Rat brain cytochromes P-450 : catalytic, immunochemical properties and inducibility of multiple forms' see the whole document -----	13,20-24

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/GB 95/02465

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
EP-A-648840	19-04-95	US-A-	5420028	30-05-95
		JP-A-	7284393	31-10-95
		EP-A-	0648841	19-04-95
		JP-A-	7298885	14-11-95
		EP-A-	0648842	19-04-95
		JP-A-	7284388	31-10-95